### (19) World Intellectual Property Organization International Bureau





#### (43) International Publication Date 9 August 2001 (09.08.2001)

#### PCT

## (10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

\_ \_ \_

(21) International Application Number: PCT/US01/03800

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) US 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560,875 (CIP)
Filed on 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

1/57188 A

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

#### NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

#### 1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

#### 2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

#### 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, \* corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

5

10

15

20

25

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

5

10

15

20

25

**30** 

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

#### 4. DETAILED DESCRIPTION OF THE INVENTION

#### **4.1 DEFINITIONS**

5

.. 10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

10

15

20

25

30

35

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4<sup>20</sup> possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match  $(1 \div 4^{25})$  times the increased probability for mismatch at each nucleotide position  $(3 \times 25)$ . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

5

10

15

20

25

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

5

10

15

20

25

30

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

5

.10

15

20

25

30

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

5

10

15

#### 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

5

10

15

20

25

30

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

5

15

20

25

30

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5

10

15

20

25

30

35

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

#### 4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

5

10

15

20

25

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

#### 4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

5

10

15

20

25

30

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

-10

. 15

20

25

30

35

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

5

10

15

20

#### 4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

5

10

15

20

25

30

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

5

10

15

20

25

30

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

5

10

15

20

25

30

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat<sup>TM</sup> kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl<sup>TM</sup> or Cibacrom blue 3GA Sepharose<sup>TM</sup>; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

5

# 4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

#### 4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered in vivo to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

## 4.9 TRANSGENIC ANIMALS

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

# 4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

#### 4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

5

10

15

20

25

30

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

# 4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

# 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells 20 include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology, J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto, 1991: deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse 25 and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 30 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

35

5

10

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

## 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

5

10

15

20

25

30

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

# 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) · as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

10

5

# 4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

20

25

30

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

5

10

15

20

25

30

35

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

# 4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on aflergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

10

15

20

25

30

35

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

5

10

15

20

25

30

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

5

10

15

20

25

30

35

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β<sub>2</sub> microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

5

10

15

20

25

30

# 4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

# 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

# 4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### 4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

5

10

15

20

25

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen'mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

# 30 4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

35

5

10

15

20

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

30

35

10

15

20

25

#### 4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

5

10

15

20

25

30

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

# 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules. that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

# 4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

#### **4.10.16 LEUKEMIAS**

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

# 4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

5

10

15

20

25

30

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

10

15

20

25

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
  - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

- (i) increased survival time of neurons in culture;
- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
  - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

# 4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape

35 (such as, for example, breast augmentation or diminution, change in bone form or shape);

5

10

15

20

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

#### 4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

# 4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

#### 4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

## **4.11.1 EXAMPLE**

5

10

15

20

25

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

# 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

5

10

15

20

25

30

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

# 4.12.1 ROUTES OF ADMINISTRATION

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

# 4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

10

15

20

25

30

35

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

5

10

15

20

25

30

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

5

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

5

10

15

20

25

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

5

10

15

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

5

10

15

### 4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC<sub>50</sub> as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD<sub>50</sub> and ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01  $\mu$ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1  $\mu$ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

### 4.12.4 PACKAGING

10

15

20

25

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

### 4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain,  $F_{ab}$ ,  $F_{ab}$  and  $F_{(ab)2}$  fragments, and an  $F_{ab}$  expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as  $IgG_1$ ,  $IgG_2$ , and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

PCT/US01/03800 WO 01/57188

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

#### 5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, 35 synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

### 5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol., 133:3001 (1984)</u>; Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

#### 5.13.2 Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

### 5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

10

15

20

25

30

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al., (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

# 5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of  $F_{ab}$  expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal  $F_{ab}$  fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an  $F_{(ab)2}$  fragment produced by pepsin digestion of an antibody molecule; (ii) an  $F_{ab}$  fragment generated by reducing the disulfide bridges of an  $F_{(ab)2}$  fragment; (iii) an  $F_{ab}$  fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv)  $F_{v}$  fragments.

## 5.13.5 Bispecific Antibodies

5

10

15

20

25

30

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

5

10

15

20

25

30

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

## 5.13.6 Heteroconjugate Antibodies

5

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

### 5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced antitumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

### 5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include <sup>212</sup>Bi, <sup>131</sup>I, <sup>131</sup>In, <sup>90</sup>Y, and <sup>186</sup>Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

### 4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

5

10

15

20

25

30

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

### 4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

5

10

15

20

25

30

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem.

5 6:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

#### 4.16 DIAGNOSTIC ASSAYS AND KITS

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

### 4.17 MEDICAL IMAGING

5

10

15

20

25

30

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

### 4.18 SCREENING ASSAYS

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
  - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

5

10

15

20

25

30

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

### 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

10

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

### 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

5

10

15

20

25

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm<sub>7</sub>, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

### 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook et al. (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

5

10

15

20

25

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (CviII\*\*), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald et al. (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a CviII\*\* digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that CviII\*\* restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

### 4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

## 5.0 EXAMPLES

#### 5.1 EXAMPLE 1

### Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

5

10

15

20

25

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

#### 5.2 EXAMPLE 2

## 5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (*i.e.*, Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from <a href="http://fasta.bioch.virginia.edu">http://fasta.bioch.virginia.edu</a>) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
	<u> </u>		976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
	]		217 225 238 271 317 404 446 469 503
	1		513-514 535 550 564 573 666-669 798
	ļ		898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
	,	ļ	147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
		•	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
		i	566 571 577 585 590 594 598 634 641
	- 1		658 666 683 725 742 764 767 786 801
		1	805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
		1	1067 1086 1088 1090 1118 1120 1122-
			1128 1142 1162 1181-1192 1199 1204
		1	1218-1219 1225 1232 1253 1267 1271-
		<del></del>	1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
			566 596 663 670 746 798 816-819 876
<del></del>	<del></del>	4727001	892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582 740 798 883 943 976 1067
preadipocytes adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
acrenai giand	Ciontech	ADRUU2	240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200
			1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
addit nout	on the same of the	1111101	118 129 132 138 151 158-163 182 195-
		1	203 215 217 238 264 269 353 384 398
	ļ	1	408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
		į	874 891 898 919 926-927 976 988
			1021 1037 1041 1062 1067 1071 1080
		-	1083 1093 1122 1131 1185 1201 1254
			1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
,			107 111 119-120 138 151 157 215 217-
		}	218 238 250 264 294 304 384 404 440
		*	446 454 477 504-505 509 514 518-519
			535 537 564 574-583 620-627 639 653
			673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
			1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
		,	446 487 564 575 844 868 910 927 976
adult lung	GIBCO	,	446 487 564 575 844 868 910 927 976 1116 8 101 111 151 187 402 446 490 514

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
	•		545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
		}	519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
		}	505 657 675 714 753 832 844 941-942
- A-14 lissen	Clareta	AT 37000	976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
		<b>[</b>	104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264 316 384 409 440 445-446 496 504 512
	•		514 518-519 535 537 549-550 564 566
	Į.		571 580 582 600 618 638 657 667 681
	ĺ		685-697 699 705 722 735-744 761 771
	·		815 833 842-865 868 875-876 918 926-
			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
			1124 1131 1144 1174 1224 1268 1331
			1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
		}	294 414 446 477 504 514 534 545 549
			592 722 873 883 952 976 1041-1042
		<u> </u>	1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
		1	238 446 497 537 642 701-706 811 877
adult bladder	7	DI DOOI	927 962 976 1083 1117 1131
aduit bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
bone marrow	Clontech	BMD001	592 607 706 873 952 1178 1329-1335 8 58-62 65-68 74 79 108 111 116 137
bolle manow	Cloniech	BIVIDUVI	147 151 164-174 213-215 238 305-307
			374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
	•	1	596 607 610 628-629 643-645 652 707-
			708 774-789 844 866-871 873 919 927
			952 963 976 998 1034 1042 1064 1083
			1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
			210 317 510-511 545 549 581 598 628
	i		638 724 766 789 844 860 868 873 919
			927 952 963 968 976 1042 1111 1141
<u></u>		4	1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	111 238 282 549 1083
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
	}	}	844 873 877 952 976 1042 1152 1268
a de de la companya della companya de la companya de la companya della companya d	Dia Cilia	- CVCV001	1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
ıı	]		335 365-367 392-401 440 466 470-471
			518 537 597 629 832 877 927 976 1006
	j		1085 1117 1129-1134 1192 1202-1205
			1 1719 1309-1378
diaphragm	BioChain	DIA002	1219 1309-1328 74 976 1083

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
***************************************	3		138 151 204-206 215-217 238 269 316
			414 433 505 510 513 550 555 580 582
			596 675 722 745 798 814 836-841 851
	}		918 976 1041 1043 1073 1083 1131
		L	1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		1
of chromosome 8	Research	ED14004	505,007
Genomic clones from the short arm	Genomic DNA from Genetic	EPM004	525 927
of chromosome 8	Research	j	
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	EFMOOD	] 331
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
Total Olam	Cionicon	LDICOG	225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
			829 836 859 909 927 943 947 963 1057
			1067-1068 1104 1135-1140 1162 1206-
		,	1207 1235 1268 1288 1307-1308 1319
			1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
			535 683 761 798 820-827 844 876 909
	<u> </u>		963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen Clontech	FKD007 FLG001	238 974 463 566 976 1074 1083 1093
fetal lung	i Ciontech	1 FLCTOUT	
Catal Issue		1	
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192
fetal lung fetal liver-spleen	Invitrogen  Columbia	1	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64
	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174
	Invitrogen  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311
fetal liver-spleen	Invitrogen  Columbia University	FLG003 FLS001	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311
	Invitrogen  Columbia University  Columbia	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311 1319 8 36-37 41-46 49 54 64 71 74 79 101
fetal liver-spleen	Invitrogen  Columbia University	FLG003 FLS001	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311 1319  8 36-37 41-46 49 54 64 71 74 79 101 111 120 129 147 207 210 215-216 238
fetal liver-spleen	Invitrogen  Columbia University  Columbia	FLG003 FLS001	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311 1319  8 36-37 41-46 49 54 64 71 74 79 101 111 120 129 147 207 210 215-216 238 250 330 353 359 366 383-384 414 478
fetal liver-spleen	Invitrogen  Columbia University  Columbia	FLG003 FLS001	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64 69-71 74 77 79 87-90 101 107 110-111 114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384 414 440 446 460 468 483 496 504-507 511-515 518-519 523 533-535 537 541 544-545 547-550 555-560 564 566 571 577 582 585-586 598 636 646-647 649 652 664 698 709-710 714 722-723 731 735-736 746-753 761 784 798 823 829 832 844 851 858-859 868 873 876 898 927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083 1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311 1319  8 36-37 41-46 49 54 64 71 74 79 101 111 120 129 147 207 210 215-216 238

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798
			851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
			1131 1144 1178 1199 1233 1240-1270
			1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
Total 11 voi	MANGO	1 L V OO1	580 722 730 749 844 918 943 976 1051
			1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
	22.1.2.084.2	11110001	425 535 537 577 598 614 836 857 1141
			1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
			225 264 316 405 422-429 488-494 496
			519 534-535 537 566 675 732 859 876-
			877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
			1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
			316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
	1		976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
			138 145 151 188 197 207 215 238 264
		· I	271 294 316 367 414 440 446 466 504
			513-514 535 542-543 550 564 571 596
			635 648-654 675 711-715 722-723 798
			832 872 876 883 927 976 1095 1144
			1168 1171 1178 1211 1335
тасгорһаде	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
	1	1	299 308-310 371-373 462 476 504 511-
	1	Ì	513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
			858-859 876 898 909 949 976 1045-
	1		1047 1076-1087 1090 1093 1116 1122
	1		1144 1209-1213 1225 1233 1256 1319
infant h		TD2002	1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
i-font basis	Columbia	IDV 6000	1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
infont h!-	University	IDC002	61 111 200 404 500 500 500 500 500 500 500 500 5
infant brain	Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
lung Charling	University	+ FD001	1221
lung , fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
	[		215 217 269 280 296 337 374-375 384
	1	1	404 446 454 475-480 498 514 518-519
			522 537 545 564 577 597 653 658 705
	1	1	721-724 754-756 779 859 868 872-874
•	!		876-877 919 927 949 951-952 959 976
	1	1	1002 1042 1048-1053 1076 1083 1088- 1089 1131 1144-1147 1216-1218 1229

m:	RNA Source	There I iharan Mana	SEQ ID NOS:
Tissue Origin	KNA Source	Hyseq Library Name	1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
lymphocytes	AICC	LFC001	634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
leukocyte	l .	LUCUUI	147 151 212 215 218 238 252 288 312-
	•		314.316.338.359.408.427.443-447.505
			510 512 514 518 534 545 549-550 561
			564 566 571 577 580 582 587-609 615
		Ĭ	632-638 658-659 698 714 725-728 832
			836 841 859 866 873-874 882-883 918-
			919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
loukocyto	Cionicon	120000	657 660 729 873 883 927 952 963 1008
			1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	Clonicon	WEEDOO!	919 929 939 952 976 1071 1118 1218
1424	ì		1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
···················· Promie			217 250-256 264 297-299 305 377-378
			398 446 481-486 505 512 537 545 549
	ł	•	571 592 725 730-733 816 829 836 844
	ł		868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
	}		1055 1076 1083 1091 1093 1116-1117
	ļ		1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
•			1319
retinoid acid induce 1	Strategene	NTR001	74 225 976
neuronal cells			
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
			1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
		<u> </u>	1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
		<u> </u>	976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
•			963 967-978 1042 1120 1152 1223-
	<u></u>		1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
			270 343-344 353 379 516 537 566 740
		1	828 927 976 979-994 1092 1153-1159
			1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
Diominon	<b>I</b>	l	995 1171
L.,			
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
thalamus			963 996-1007 1059 1093 1160-1162
L.,	Clonetech	THA002 THM001	963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639
thalamus			963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639 750 867 874 878-881 927 963 1023
thalamus	Clonetech	THM001	963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639 750 867 874 878-881 927 963 1023 1083 1094-1096 1124
thalamus			963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639 750 867 874 878-881 927 963 1023 1083 1094-1096 1124 8 61 114 129 132 210 225 231 306
thalamus	Clonetech	THM001	963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639 750 867 874 878-881 927 963 1023 1083 1094-1096 1124 8 61 114 129 132 210 225 231 306 317-319 336 340 359 380 398 446 448-
thalamus	Clonetech	THM001	963 996-1007 1059 1093 1160-1162 8 120 151 208 221 316-317 353 639 750 867 874 878-881 927 963 1023 1083 1094-1096 1124 8 61 114 129 132 210 225 231 306

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
		<u> </u>	1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
			320-326 345-352 361 381-382 446 467
			483 514 534 549-550 564 578 602 649
	1		844 882-883 927 950 956 976 1008-
Ì		}	1028 1076 1083 1117-1120 1142 1163-
<u></u>	<u> </u>		1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
1			545 592 611 873 883-884 927
1			952 1029-1031 1042 1151-1152
	<u> </u>		1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
1			885-886 976 1001 1032-1033
		<u> </u>	1232

TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:	004067	<del>  ,,                                  </del>	Human secreted protein, SEQ ID NO: 8148.	83	42
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
31	G03371	Homo sapiens		58	32
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	2457	98
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.		95
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
<del>42</del> 43	X61048	Hydra sp.	mini-collagen	105	35
43 44	M76546	Helianthus	hydroxyproline-rich protein	110	31
44	U82288	annuus Caenorhabditi	Rac-like GTPase	139	70
		s elegans		118	58
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	113	63
47	AF090942	Homo sapiens	PRO0657	90	59
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	72	56
49	AJ005560	Mus musculus	SPR2B protein		98
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63	U35376	Homo sapiens	repressor transcriptional factor	485	65
64	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	785	74
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
60	C02065	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
69	G02965		Human oxidoreductase YTFO3.	1144	98
70	W75770	Homo sapiens	!	239	76
71 72	AB011135 AB014885	Homo sapiens Halocynthia	KIAA0563 protein HrPOPK-1	813	78
	1	roretzi	phospholipase B	955	73
73	AF045454	Cavia porcellus	phospholipase B	933	1'3

SEQ	Accession	Species	Description	Smith-	%
ID	No.	[		Waterman	Identity
NO:	<del> </del>	musculus		Score	<del> </del>
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
,,	100020	norvegicus	gp210 (AA 1-1660)	413	04
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	351	54
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha-	1357	99
			1-I (hCavT3).		
79	Y14591	Human papillomaviru s type 68	APM-1 protein	767	100
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis thaliana	protein arginine N-methyltransferase-like protein	359	65
82	L46815	Mus musculus	DNA binding protein Rc	895	75
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	538	71
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid sequence SEQ ID NO:100.	156	48
88	AJ225124	Mus musculus	hyperpolarization-activated cation channel, HAC3 -	487	95
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93 94	AF170723 X13292	Homo sapiens Trypanosoma	protein kinase STK10	401	53
94 .	X13292	brucei	GPI-phospholipase C (AA 1 - 358)	151	37
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus norvegicus	sodium channel protein I (aa 1-2009)	1775	92
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	675	48
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence difference at residue 58	160	60
102	U22829	Mus musculus	P2Y purinoceptor	264	42
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled receptor-B3.	516	99
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119.	343	57
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia coli	sialic acid transporter	587	96
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	182	94
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
112	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84.	274	51
113	AF016365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AC007956	Homo sapiens	unknown	520	75
115	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets hornologous factor (ets- domain transcription factor ESE-3A, isoform 1))	484	91
117	W18084	Homo sapiens	Human Aurora-2.	546	87

				0.34	T 0/
SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
		norvegicus			
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr	1646	94
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC	373	68
121	03/3/2	Tiomo supiens	3.1.3.48}	- ''	50
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi	contains similarity to C2 domains	219	29
124	000107	s elegans			
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
	ļ	•	4	ļ	1
126	AB021861	Mus	apoptosis signal-regulating kinase 2	153	65
	1	musculus			
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter	807	97
	ļ		hCNT3	Í	<u> </u>
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF201734	Mus	testis specific serine kinase-3	800	87
		musculus	•	1	1
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
			clone HSQEL25.		
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	5041	99
		•	similarity to P49205 (PID:g1345860)	i	1
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis	DNA-damage inducible protein DDI1-like	147	55
		thaliana			
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus	AMPA receptor binding protein	623	93
		norvegicus		ļ	
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus	transmembrane receptor UNC5H2	578	84
}		norvegicus			J
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein	727	92
	<u> </u>		M160 precursor	<u> </u>	_L
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia	HrsA	818	90
L	<u></u>	coli			
149	M83316	Escherichia	pppGpp phosphohydrolase	915	95
	<u> </u>	coli			
150	AL163279	Homo sapiens	homolog to cAMP response element binding and	1261	99
		<u> </u>	beta transducin family proteins		<del> </del>
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	392	61
		<u> </u>	ligand (clone 3TW).		<del> </del>
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus	zinc finger protein	352	74
	<del> </del>	musculus		535	1
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual	670	98
1	1,40,000	1	specifity Ser/Thr/Tyr kinase domain	\	<del> </del>
160	AJ250425	Rattus	Collybistin I	556	74
121	C02005	norvegicus	11	270	100
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	1100

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	] •	•	Waterman	Identity
NO:	1			Score	
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus	calcium transporter CaT1	700	96
105		norvegicus	•	/00	1 30
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
171	AF035268	gen. sp.  Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus	semaphorin cytoplasmic domain-associated	507	82
173	Y27918	musculus Homo sapiens	Protein 3B Human secreted protein encoded by gene No.	653	99
	1	1	123.		
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
		<u> </u>	codon)		
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens.	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria	casein kinase II alpha subunit	364	50
104	1	parva	Manhara Land : BBC:0:0	146	100
194 195	Y66645 W95631	Homo sapiens Homo sapiens	Membrane-bound protein PRO1310.  Homo sapiens secreted protein gene clone	382	90 49
196	AF255614	Rattus	hj968_2. scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
		thaliana			
198	AF073967	Mus musculus domesticus	olfactory receptor	316	43
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
	1 33 46 / / 40	TTOTALO SUPTORIS	DPH2L=candidate tumor suppressor gene	102	

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	<b> </b>
	<u> </u>		{ovarian cancer critical region of deletion}		<u> </u>
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific protein, calphotin.	541	82
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus norvegicus	ankyrin binding cell adhesion molecule neurofascin	471	69
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus musculus	FYVE finger-containing phosphoinositide kinase	933	93
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein	563	78
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus musculus	Kupffer cell receptor	567	40
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	636	96
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus	•	ļ	
226	AE000218	Escherichia coli	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus musculus	GTP-binding like protein 2	265	88
229	AF122924	Xenopus laevis	Wnt inhibitory factor-1	316	40
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific phospholipase-D.	290	100
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus musculus	Embryo Brain Kinase	460	62
238	U64857	Caenorhabditi s elegans	similar to the BPTI/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor (TFPI)	284	33
239	AJ250840	Mus musculus	serine/threonine protein kinase	739	63
240	AJ223472	Mus musculus	transcription elongation factor TFIIS.h	222	38
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein sequence SEQ ID NO:18.	353	52
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus norvegicus	orphan transporter v7-3	667	93
244	AF016191	Rattus norvegicus	potassium channel	1043	98
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	645	98
245	Y29868	Homo sapiens	Human secreted protein clone pp325 9.	497	98
47U		Homo sapiens	Not4-Np	188	83
	I AHIXIIA''S		1 11017-170	1 100	1 00
247 248	AF180475 Y17227	Homo sapiens	Human secreted protein (clone yal-1).	690	99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
	<del> </del>	sexta	protein SCLP		┼──
250	AF192756	Kaposi's sarcoma- associated herpesvirus	Orf73	134	34
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus musculus	DNA binding protein Rc	251	67
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus musculus	Citron-K kinase	1201	98
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus cuniculus	Phospholipase	368	80
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus musculus	L-periaxin	430	72
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus norvegicus	SLIT-2	198	40
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor (GPCR).	636	99
266	U27269	Mus musculus	sodium glucose cotransporter	204	56
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus norvegicus	putative taste receptor TR1	209	39
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus pyogenes	Fc-gamma receptor	129	26
271	AB009883	Nicotiana tabacum	KED	109	26
272	AF137367	Mus musculus	VPS10 domain receptor protein SORCS	899	97
273	L34938	Rattus norvegicus	ionotropic glutamate receptor	460	86
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)	188	74
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	173	94
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis thaliana	Contains PF 00069 Eukaryotic protein kinase domain.	157	43
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
283	AF156530	Mus musculus	ETS-domain transcriptional repressor PE1	605	76
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate reading frame protein.	647	100
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26.	300	90
286	AF016411	Homo sapiens	KCNA3.1B	137	100
287	W89253	Homo sapiens	Human ALP.	688	97
288	AF112886	Bos taurus	differentiation enhancing factor 1	750	96
289	AF113131	Homo sapiens	host cell factor homolog LCP	367	44
290	U52111	Homo sapiens	plexin-related protein	698	100
291	AF026504	Rattus	SPA-1 like protein p1294	603	89

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Bestription	Waterman	Identity
NO:	110.	1		Score	1001111
NO.	<del> </del>	norvegicus		50010	<del> </del> -
292	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
292	AL 102054	norvegicus	interacting protein 2 Maguin-2	124	1 33
293	X99211	Drosophila	ubiquitin-specific protease	143	38
293	733211	melanogaster	doiquidit-specific protease	143	1 30
294	Y94943	Homo sapiens	Human secreted protein clone yt14 1 protein	185	94
	1,7,7,0	The section of the se	sequence SEQ ID NO:92.		
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577 1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14 1 protein	182	97
270	17.7.5	Trome suprems	sequence SEQ ID NO:92.	.02	1
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	605	69
227	1 200300	120,200	gene 16 SEQ ID NO:63.		"
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
303	043380	Homo sapicits	Method: conceptual translation supplied by	720	1 '2
		}	author		
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
307	D16512	musculus	membrane grycoprotem	199	**
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
300	AF255014	norvegicus	scarrolding protein SELFR	039	00
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
311	003413	discoideum	calcium binding protein	131	30
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
312	10/34/	Tionio sapiciis	124 SEQ ID NO:124.	'**	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
314	AC004010	Tromo sapicus	44% similarity to U42767 (PID:g1736918)	.,,	130
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and	278	38
313	711021372	1101110 Supiciis	GENEWISE)	270	30
316	U70209	Mus	polycystic kidney disease I protein	165	38
2.0	570205	musculus	polyeyone maney allocate i protein		1 55
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
01,	127070.5	norvegicus	oonsaard adda		50
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma	activated protein kinase C receptor homolog	141	38
		vivax			1
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
	1 =====	Septemb	receptor		1
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
- <del></del>		1	123.	1	1
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214	97
3_0		Tromo suprems	3.1.4.37)		-
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
	X75756	Homo sapiens	protein kinase C mu	540	78
327		Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
327 328	G02292	friction outsiding			99
328	G02292 AF168990		l billative († 1 P-binding brotein	1 877	
328 329	AF168990	Homo sapiens	putative GTP-binding protein	877 581	
328			anti-HIV gp120 antibody heavy chain variable	581	80
328 329 330	AF168990 S67984	Homo sapiens Homo sapiens	anti-HIV gp120 antibody heavy chain variable region	581	80
328 329 330 331	AF168990 S67984 X13916	Homo sapiens Homo sapiens Homo sapiens	anti-HIV gp120 antibody heavy chain variable region  LDL-receptor related precursor (AA -19 to 4525)	581 2823	98
328 329 330	AF168990 S67984	Homo sapiens Homo sapiens	anti-HIV gp120 antibody heavy chain variable region  LDL-receptor related precursor (AA -19 to 4525)  Human signal peptide containing protein HSPP-	581	80
328 329 330 331	AF168990 S67984 X13916	Homo sapiens Homo sapiens Homo sapiens	anti-HIV gp120 antibody heavy chain variable region  LDL-receptor related precursor (AA -19 to 4525)	581 2823	98

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	I	<u> </u>		Score	<u>↓</u>
	<u> </u>		similarity to P49205 (PID:g1345860)		<u> </u>
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	1111	67
	<u> </u>	<u> </u>	124 SEQ ID NO:124.		
336	AF006466	Mus	lymphocyte specific formin related protein	193	75
	<u> </u>	musculus	•		ì
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	632	97
	1	1	APOLLON		1
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia	L-idonate transcriptional regulator	928	98
J	1.2000.57	coli	O location automorphism regulator	720	1 20
342	D90855	Escherichia	glycerol-3-phosphate dehydrogenase (EC	769	99
J-12	250055	coli	1.1.99.5) chain A, anaerobic	702	"
343	D85613	Escherichia	membrane component	399	100
J4J	D03013	coli	Memorane component	377	100
344	M93239	Escherichia	transmembrane protein	232	100
3 <del>44</del>	M193239		transmemorane protein	232	100
345	M60177	coli   Escherichia		750	100
<b>34</b> 3	MOU1//		enterobactin	759	99
246	D00400	coli		(30	<del> </del>
346	D90699	Escherichia	Sensor protein copS (EC 2.7.3).	638	97
		coli			
347	D90843	Escherichia	CapB protein.	552	100
	<u> </u>	coli			<u> </u>
348	M13422	Escherichia	49 kd protein	1193	96
		coli			
349	L10328	Escherichia	similar to drug resistance translocases	340	90
	1	coli			1
350	X69942	Mus	enhancer-trap-locus-1	560	82
		musculus	-		
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
		1	activated potassium channel		1
352	D90777	Escherichia	3-hydroxybutyryl-CoA dehydrogenase (EC	577	100
	į	coli	1.1.1.157) (b- hydroxybutyryl-CoA		)
			dehydrogenase) (BhbD).		
353	D90863	Escherichia	similar to	311	98
		coli			1
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-	482	55
	1		7).		
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVHI	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID	165	100
550	10/21/	Tionio sapiens	NO:258.	103	100
				233	93
250	100133	Uama conienc			
	J00132	Homo sapiens	beta-fibrinogen		
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
360 361	G03789 R28916	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).	128 108	70 40
360 361	G03789	Homo sapiens Homo sapiens Rattus	Human secreted protein, SEQ ID NO: 7870.	128	70
360 361 362	G03789 R28916 U16655	Homo sapiens Homo sapiens Rattus norvegicus	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4	128 108 649	70 40 65
360 361 362 363	G03789 R28916 U16655 G03119	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.	128 108 649	70 40 65 42
360 361 362 363 364	G03789 R28916 U16655 G03119 U47276	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2	128 108 649 95 104	70 40 65 42 34
360 361 362 363 364 365	G03789 R28916 U16655 G03119 U47276 G03789	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.	128 108 649 95 104 183	70 40 65 42 34 65
360 361 362 363 364 365 366	G03789 R28916 U16655 G03119 U47276 G03789 G04091	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172.	128 108 649 95 104 183 118	70 40 65 42 34
360 361 362 363 364 365 366 367	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172.  M-phase phosphoprotein 9	128 108 649 95 104 183	70 40 65 42 34 65
360 361 362 363 364 365 366 367	G03789 R28916 U16655 G03119 U47276 G03789 G04091	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172.	128 108 649 95 104 183 118	70 40 65 42 34 65 46
360 361 362 363 364 365 366 367 368	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200.  chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172.  M-phase phosphoprotein 9	128 108 649 95 104 183 118 564	70 40 65 42 34 65 46 75
359 360 361 362 363 364 365 366 367 368 369	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 cICK0721Q.3 (Kinesin related protein)	128 108 649 95 104 183 118 564 3387	70 40 65 42 34 65 46 75 99
360 361 362 363 364 365 366 367 368	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 cICK0721Q.3 (Kinesin related protein) reverse transcriptase	128 108 649 95 104 183 118 564 3387 92	70 40 65 42 34 65 46 75 99 59
360 361 362 363 364 365 366 367 368 369	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 cICK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase	128 108 649 95 104 183 118 564 3387	70 40 65 42 34 65 46 75 99
360 361 362 363 364 365 366 367 368 369 370	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 cICK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase like	128 108 649 95 104 183 118 564 3387 92	70 40 65 42 34 65 46 75 99 59
360 361 362 363 364 365 366 367 368 369 370	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932 X86400	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870.  Human secreted protein, SEQ ID NO: 8172.  M-phase phosphoprotein 9 cICK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase like  Human secreted protein, SEQ ID NO: 7253.	128 108 649 95 104 183 118 564 3387 92 242	70 40 65 42 34 65 46 75 99 59
360 361 362 363 364 365 366 367 368 369 370 371 372	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932 X86400 G03172 U49974	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 clCK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase like Human secreted protein, SEQ ID NO: 7253. mariner transposase	128 108 649 95 104 183 118 564 3387 92 242 165 257	70 40 65 42 34 65 46 75 99 59 73
360 361 362 363 364 365 366 367 368 369 370 371 372 373	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932 X86400 G03172 U49974 X13916	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 clCK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase like Human secreted protein, SEQ ID NO: 7253. mariner transposase  LDL-receptor related precursor (AA -19 to 4525)	128 108 649 95 104 183 118 564 3387 92 242 165 257 21193	70 40 65 42 34 65 46 75 99 59 73
360 361 362 363 364 365 366 367 368 369	G03789 R28916 U16655 G03119 U47276 G03789 G04091 X98258 AL021366 U70932 X86400 G03172 U49974	Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Gallus gallus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Peromyscus leucopus Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.  Type III procollagen (prior art).  phospholipase C delta-4  Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2  Human secreted protein, SEQ ID NO: 7870. Human secreted protein, SEQ ID NO: 8172. M-phase phosphoprotein 9 clCK0721Q.3 (Kinesin related protein) reverse transcriptase  gamma subunit of sodium potassium ATPase like Human secreted protein, SEQ ID NO: 7253. mariner transposase	128 108 649 95 104 183 118 564 3387 92 242 165 257	70 40 65 42 34 65 46 75 99 59 73

SEQ	Accession	Species	Description	Smith-	1%
ID SEQ	No.	) Species	Description	Waterman	Identity
NO:	1.0.	•		Score	lacuity
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus Mus	GTP binding protein	1456	91
576	7132374	musculus	G11 bliddig plotem	1450	122
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium	protein tyrosine kinase	115	44
302	001030	discoideum	process cyrosine katabo	115	1
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by	171	34
	1 /0332	riomo sapiens	gene 38.	1/1	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates	dopamine receptor D4	105	35
	ì	syndactylus		}	1
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus	ADP-ribosylation factor-directed GTPase	545	89
		musculus	activating protein isoform b	L	<u> </u>
405	X92887	Human	pol/env	162	30
	]	endogenous		]	
106	1100160	retrovirus K		L	L
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone	2004	99
!	}	•	HTSEV09.		
411	AB043953	Mus	Chat-H	2628	82
		musculus		<u> </u>	<u> </u>
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID	1014	92
413	U10542	Pan	NO:148. MHC class I A	265	71
413	010542		MHC class I A	263	/1
414	AF155097	troglodytes	NN/ DENI 7 anti-an	950	05
		Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar	1446	96
	⊥	<u> </u>	to rat sperm antigen 4 (SPAG4)))	L	
100			F24J5.4	112	35
422	AC008075	Arabidopsis			

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ	1961	99
423	133342	1101110 Sapiciis	ID NO. 191.	1901	1 33
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21 1, SEQ ID NO:20.	768	98
<del>428</del> 429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
		<b>.</b>	associated protein.		
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon esculentum	extensin-like protein	613	48
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus norvegicus	transmembrane receptor UNCSH1	817	93
450	AF081249	Homo sapiens	TANKI placed matrix MONITIA languagement	4568	<del> </del>
451	AC005498		JAW1-related protein MRVI1A long isoform R31665 1		99
452	M60235	Homo sapiens Homo sapiens	granule membrane protein-140	316 464	62
453	AB036706	Homo sapiens	intelectin	730	73 88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
456	Y36705	Homo sapiens	(CIRP-1). Fragment of human secreted protein encoded by	106	40
457	N91325	Homo sapiens	gene 62.	1 2202	100
457			DNA encoding human growth hormone receptor.	3282	96
4.20	M19155	Plasmodium falciparum	S-antigen precursor	110	36
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
460	Y14482		clone HTDAD22.		
		Homo sapiens	Fragment of human secreted protein encoded by gene 17.	184	54
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein sequence SEQ ID NO:16.	135	47
463	X84960	Triticum aestivum	low molecular weight glutenin	109	33
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	-85
465	AF189764	Mus musculus	alpha/beta hydrolase-1	502	59
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by gene 77.	1172	99
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	Pl3-kinase	5832	97
470	X70922	Mus musculus	neurotoxin homologue	118	47
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	72	13

SEQ	Accession	Species	Description	Smith-	%
ID	No.	] -	•	Waterman	Identity
NO:				Score	1
	<del> </del>	†	gene 62.	<del> </del>	<del>                                     </del>
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
7/7	10,007	Tromo sapions	sequence.	1015	1 "
475	W93254	Homo sapiens	Human ESRPI protein.	943	80
					1
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	202	60
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777			3427	92
419	AF102///	Mus musculus	FYVE finger-containing phosphoinositide kinase	3427	92
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
701	1,101,101	Tiomo sapiens	SYTAX1.		''
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
700	10151/4		3	147	,3
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
	1			197	66
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.		
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein 3	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	197	67
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
777	710000257	familiaris	l	1 20	1 40
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935			213	52
477	070933	Peromyscus	reverse transcriptase	213	32
500	1140500	maniculatus		25406	
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135_9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
	1	'	NO:180.	L''.	
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
509	U43360	Peromyscus	polypeptide kinase))	97	62
JUT	U4330V	maniculatus	reverse transcriptase	"	02
510	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
511	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
513	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila melanogaster	CG6393 gene product	259	42 .
515	Z17206	Xenopus	p46XIEg22	128	40
	1	laevis			L
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
517			HSPC249	444	98
	AF151083	I HOMO Saniene			
517 518 ·	AF151083	Homo sapiens			
	AF151083 S80864 X92485	Homo sapiens Homo sapiens Plasmodium	cytochrome c-like polypeptide	318	50

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.	opocios	Description	Waterman	Identity
NO:	1	1		Score	Identity
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857	Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	253	73
324	1100027	110mo sapiens	HPMBQ32.	253	/3
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528	U47924	Homo sapiens	C8	1112	86
529	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	gin-induced kinase		
538	AF135022	Homo sapiens	mediator	206	53
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	128	100 59
540	AF016430	Caenorhabditi	contains similarity to a BR-C/TTK domain	141	
340	AF010430	s elegans	contains similarity to a BR-C/11K domain	853	39
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45%	408	1
J41	AC003033	Homo Sapiens	similarity to P22059 (PID:g129308)	408	66
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	101
543	AF102530	Mus	olfactory receptor F3	327	73
243	A1 102330	musculus	Onactory receptor 13	321	13
544	Y73431	Homo sapiens	Human secreted protein clone yb186 1 protein	386	100
247	175451	Tionio sapiens	sequence SEQ ID NO:84.	300	100
545	AE004833	Pseudomonas	probable TonB-dependent receptor	279	42
343	7111004055	aeruginosa	probable rollb-dependent receptor	213	42
546	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein	1772	67
	103132	Tromo suprems	B receptor protein.	1772	1 67
548	Y91493	Homo sapiens	Human secreted protein sequence encoded by	176	100
		Tromb baprens	gene 43 SEQ ID NO:166.	170	100
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584 2 protein	1224	94
			sequence.	1224	74
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
554	AB025258	Mus	granuphilin-a	501	41
		musculus	, ————————————————————————————————————	1	**
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein	183	32
	1		complex component TRAP150	l	
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus	Ca2+ dependent activator protein for secretion	1010	93
	1	musculus	P		"
561	AF187325	Canis	melanoma antigen	287	55
	}	familiaris		J	
562	AJ001981	Homo sapiens	OXAIL	2512	99
563	Z17238	Rattus	glutarnate receptor subtype delta-1	338	66
		norvegicus	, O		1
564	W30638	Homo sapiens	Partial human 7-transmembrane receptor	371	100
	1		HAPO167 protein.	)	1
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid	1138	78
_	+		sequence SEQ ID NO:63.	<b>-</b>	1
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
568	AF151043	Homo sapiens	HSPC209	798	100
	<del></del>		L		

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	1.	•	Waterman	Identity
NO:		}		Score	1
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
<del>571</del>	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
3/1	107030	110mo sapions	sequence.	1004	1 200
572	AT 021177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
	AL031177			254	45
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.		
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
	{	1	ID NO:388.		
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
202	AB030234	familiaris	D7 dopamine receptor	- T	1 30
504	GOOGL		Human peareted pretain CEO ID MO. 6700	345	90
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.		85
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
	1,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	<del></del>	Antigen)		L
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
		musculus			
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	81
		_	HPMBQ32.		
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
			protein.		
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
3,72	1330,3	TIOING Supions	polypeptide.	150)	1
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	97
J <b>J</b> J	1 23031	radillo sapiens	sequence SEQ ID NO:108.	1112	),
594	Y27658		Human secreted protein encoded by gene No. 92.	763	79
		Homo sapiens			58
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	
596	AF151110	Mus	COP1 protein	2215	95
		musculus			<u> </u>
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598.	AF192499	Mus	putative secreted protein ZSIG37	143	40
		musculus			<u> </u>
599	AF119855	Homo sapiens	PRO1847	236	76
600	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
004	AL070020	Homo sapiciis	Antigen)	1333	"
605	AD022106	+ TY	KIAA1280 protein	2015	100
605	AB033106	Homo sapiens		3915	
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
	1	1	HPMBQ32.		
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
		1	107.		
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
013	1 02033	rionio sapiens	l	173	1 39
<u> </u>	1407052	Potture.	clone HTDAD22.	450	04
614	M87053	Rattus	lens membrane protein	450	84
	1	norvegicus			<u> </u>
615	AC004232	Homo sapiens	FPM315	163	37
616	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein of clone CS/32-3.  Human secreted protein vb12 1, SEQ ID NO:4.	733	
624	AF034745	Mus	LNXp80		100
		musculus		637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
547	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
651	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
			111 NHQ32.		1
658 659	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134	65

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	375	43
662			designated HSCOP-6.		<u> </u>
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein (hs6M1-21))	480	55
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KIAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
			ligand (clone 2DD).		1
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709 710	Y96202 M63577	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
		Saccharomyc es cerevisiae	SFP1	131	59
711	AB026291	Rattus norvegicus	acetoacetyl-CoA synthetase	467	85
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota marmota	olfactory receptor	615	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid receptor beta4 subunit	578	99
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	570	74
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma tigrinum	electrogenic Na+ bicarbonate cotransporter; NBC	111	41
724	AF127084	Mus musculus	semaphorin cytoplasmic domain-associated protein 3A	5253	94
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus norvegicus	potassium channel	370	100
727	AB029559	Rattus .	BATI	139	35
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila) homolog)	142	68
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor homologue Vanilrep1.	675	99
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia coli	putative transport protein	592	97
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor (rhodopsin family) protein similar to high- affinity lysophosphatidic acid receptor homolog)	2173	99
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99.
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus musculus	open reading frame (196 AA)	83	24
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32 clone HLTCJ63.	448	95
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA from plasmid pGCS2232.	968	65
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein sequence SEQ ID NO:76.	622	100
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	391	87

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus	granuphilin-a	773	41
		musculus	***	000	99
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus	vasopressin receptor	979	68
		norvegicus			<u></u>
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
761		_1		4433	99
762	AB020676	Homo sapiens	KIAA0869 protein	2285	99
763	AK026992	Homo sapiens	unnamed protein product		
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774 774	AB025258	Mus	granuphilin-a	500	41
	1	musculus	l		93
775	AF125101	Homo sapiens	HSPC040 protein	232	
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434 .	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOVI	1904	91
786	Y91870	Homo sapiens		547	100
			Human apoptosis related protein.	1062	94
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.		98
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
	Y69384	Homo sapiens	Amino acid sequence of a 14274 receptor	119	100
795	ì	1	protein.	I	i i

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	A D259240	770	,	Score	
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CRI protein.	11963	97
	<u> </u>	(human)	<u></u>	Ì	1
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172 1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
			LAT2	1304	1 30
809	W70321	Homo sapiens	Secreted protein CC198 1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
	1		clone HOVBA03.	1	1
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
015	74 203772	Tiome sapiens	encoded by GenBank Accession Number L25899	/64	100
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone gnl 14 1.	358	100
817	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
818	AF151800	Homo sapiens	CGI-41 protein	1106	95
819	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
820	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220			1	
823	Y50125	Homo sapiens	TERA	396	48
ديه	1 30123	Homo sapiens	Human glycophosphatidylinositol-anchored protein GPI-122.	4897	99
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma- 2 subunit	1105	100
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded from gene 28.	1540	100
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
		<u></u>	gene 24 SEQ ID NO:147.	541	
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262 .	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828 1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi s elegans	glycine-rich	85	36
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	998	75
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543 3.	1105	67
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens		644	
842			Human secreted protein, SEQ ID NO: 6731.		97
	AF036717	Homo sapiens	FGFR signalling adaptor SNT-1	2629	99
843	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	1089	100
844	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	357	69
845	AF151810	Homo sapiens	CGI-52 protein	1443	88
846	X83378	Homo sapiens	putative chloride channel	1620	99
847	AC004883	Homo sapiens	similar to general transcription factor 21; similar	655	96

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:			<u> </u>	Score	
			to AF038969 (PID:g2827207)		
848	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
850	4 D02022	177	P34984 (PID:g464305)	1000	1.00
851	AB038237 AF124490	Homo sapiens	G protein-coupled receptor C5L2	1767	100
852	Y86217	Homo sapiens	ARF GTPase-activating protein GIT1 Human secreted protein HWHGU54, SEQ ID	3415	98
		Homo sapiens	NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (	1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1048	98
		_	gene 25 SEQ ID NO:305.		-
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 ·	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus norvegicus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
885	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887	X92744	Homo sapiens	hBD-1	375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone cn621_8.	994	94
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595	99
890	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo sapiens	G0S19-2 peptide precursor	481	100
897	Z68747	Homo sapiens	imogen 38	2018	99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
899	AF225420	Homo sapiens	AD025	734	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1		Waterman	Identity
NO:	DCOCET	1,7		Score	1.00-
900	P60657 M27288	Homo sapiens	Sequence of human lipocortin.	1835	100
901	W85737	Homo sapiens	oncostatin M	1297	100
902	G01349	Homo sapiens	Polypeptide with transmembrane domain.	749 650	99
903	Y00261	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 5430.		99
904 905	AF039688	Homo sapiens	Human secreted protein encoded by gene 4. antigen NY-CO-3	771	99
	AB007836				1
906 907	AB007836 AB017507	Homo sapiens	Hic-5	2544 224	100
	AK000056	Homo sapiens Homo sapiens	Apg12		100
908	Y86299		unnamed protein product	1537	98
909	1 80299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162 1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48 1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
930	A30363	familiaris	1406	1004	100
937	B08906	Homo sapiens	Human secreted protein sequence encoded by	117	44
028	N112602	Tramp	gene 16 SEQ ID NO:63.	1000	100
938	M13692 Y53886	Homo sapiens Homo sapiens	alpha-1 acid glycoprotein precursor  A suppressor of cytokine signalling protein	1064 515	99 42
940	Y16630	Horno sapiens	designated HSCOP-6.  Human Putative Adrenomedullin Receptor	1904	99
941	AC005102	Homo sapiens	(PAR). small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ	667	100
	<u> </u>		ID NO. 463.	<u> </u>	
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEO	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
		•	ID NO. 496.		i
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949 3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202 3 protein	587	100
,,,	133032	Tiomo suprons	sequence SEQ ID NO:110.	307	100
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc#	1466	100
303	AB020313	rionio sapiens	AF030433	1400	100
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
913	AJ132429	nomo sapiens	gated cation channel hHCN4	0293	100
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens		2552	100
978	AF117657		cystine/glutamate transporter thyroid hormone receptor-associated protein	3348	99
9/6	AF11/05/	Homo sapiens	complex component TRAP80	3346	199
979	AF044201	Rattus	neural membrane protein 35; NMP35	1570	92
717	AF044201	norvegicus	nediai memorane protein 55, MMF55	1370	1 32
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-	1553	99
702	W 60423	Homo sapiens	encoded protein.	1555	)
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	172	70
703	1 14402	Tromo sapiens	gene 17.	11/2	/3
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca	juvenile hormone esterase binding protein	226	32
200	T 122420	sexta	Javenne normone esterase omdnig brotem	1220	32
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
990	AF204159	Homo sapiens	potassium large conductance calcium-activated	1486	100
330	A1204133	110mo sapiens	channel beta 3a subunit	1400	100
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi	VM106R.1	327	40
174	AL031200	s elegans	AMITOURI	1 321	1 ***
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	199
993	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
				5811	99
995	AF133845	Homo sapiens	thursid harmone recenter associated protein	4999	100
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein	4777	100
007	17/62066	II	complex component TRAP150	1204	102
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID	725	100
000	3712220	+ TT	NO:212.	1051	1-00
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vB_1, SEQ ID NO:56.  membrane associated protein SLP-2	676 1747	100
1001	AF190167	Homo sapiens			

SEQ	Accession	Species	Description	Smith-	%
ID	No.		-	Waterman	Identity
NO:				Score	
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No.	2150	100
			24.	2130	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor		98
1007	Y35997			326	
		Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010 ·	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.		
1017	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	1876	100
				589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMP1	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.		99
1031	AK023007	Homo sapiens	unnamed protein product	1354	
1032	-!		unianed protein product	766	100
	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1042	G00308	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.  Human secreted protein, SEQ ID NO: 6613.		
1043	M94582	Homo sapiens	interleukin 8 receptor B	60	53
1044				1850	100
1043	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	500	100
1046				580	100
	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049 ·	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	¥27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	100
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	2014	99
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	148	50
1071	X03145	Homo sapiens	pot. ORF III dJ889M15.3 (novel protein)	821	91
1072	AL031177	Homo sapiens	gpStaf50	249	62
1073	X82200	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1074 1075	G03213 Y36233	Homo sapiens Homo sapiens	Human secreted protein, SEQ 15 NO. 7294.  Human secreted protein encoded by gene 10.	506	55
1075	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084	G03564	Homo sapiens	Hurnan secreted protein, SEQ ID NO: 7645.	83	51
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063 Y02693	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.  Human secreted protein encoded by gene 44	83 149	35 59
1100		Homo sapiens			

SEQ ID	Accession No.	Species	Description	Smith- Waterman	%
NO:	No.			Score	Identity
1101	AF119851	Homo sapiens	PRO1722	L	<del> </del>
	<del></del>			183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus musculus	ribosomal protein L28	128	69
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by		
		1	gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 155.	244	97
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469 ·	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by gene 49.SEO ID NO:170.	542	100
1134	AB017908	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis familiaris	D4 dopamine receptor	89	48
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7000.	512	85
	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1148					

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Bescription	Waterman	Identity
NO:	1	1		Score	100
110.	<del> </del>	s elegans	cerevisiae zinc resistance protein	1	<del> </del>
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181,	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
1154	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens	Tumour suppressor protein, p53.	341	87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	99	41
		<b> </b>	NO:180.		
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	149	51
			gene 17.		
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocysteine associated protein	249	62
	1 - 1 - 1 - 1	norvegicus			
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22	107	71
	<u> </u>		clone HMSJW18.		<u> </u>
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens CI542 2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID	1089	89
			NO:175.		<u> </u>
1200	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	154	57

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	- P		Waterman	Identity
NO:				Score	Identity
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein		
		1	sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.  Human secreted protein, SEQ ID NO: 5090.	351	
1213	AF090942	Homo sapiens	PRO0657	124	73
1214	Y14427	Homo sapiens			70
			Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	2) polypeptide.  High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB).  Extended human secreted protein sequence, SEQ	135	31
1223	Y00278	Homo sapiens	ID NO. 160.	-	<u> </u>
1224	AF161422	Homo sapiens	Human secreted protein encoded by gene 21. HSPC304	260	95
1225	U14970			568	90
1225	G01733	Homo sapiens	ribosomal protein S5	202	95
1227	AF099973	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
		Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YPIB	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens		607	100
1240	AF004161	Oryctolagus	Homo sapiens CH27_1 clone secreted protein.	697	98
	İ	cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
		Homo sapiens			39

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:				Score	<b>-</b>
10.40	AF072509	Rattus	ID NO. 160. glutamate receptor interacting protein 2	559	90
1248	AF0/2509	norvegicus	glutamate receptor interacting protein 2	339	1 30
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	<u> </u>		gene 27 SEQ ID NO:131.	<u> </u>	<u> </u>
1251	L15313	Caenorhabditi s elegans	putative	858	59
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate reading frame protein.	278	75
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-l	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
1260	R95332	Homo sapiens	from gene 26. Tumor necrosis factor receptor 1 death domain	986	100
			ligand (clone 3TW).		
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor sequence.	288	71
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ ID NO:2.	723	93
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus norvegicus	phosphatidylinositol 5-phosphate 4-kinase gamma	859	95
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus musculus	LMBR2	552	76
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras GTPase-activating protein p135 SynGAP)	820	98
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthethase (acetate-coA ligase))	1280	100
1274	AF064748	Mus musculus	S3-12	3523	61
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted protein.	643	90
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus cuniculus	aorta CNG channel (rACNG)	267	85
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus musculus	odd-skipped related 1 protein	357	98
1284	U87318	Xenopus laevis	NaDC-2	535	60
1285	AF061346	Mus musculus	Edp1 protein	452	68
1286	AB030182	Mus musculus	contains transmembrane (TM) region	582	68
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
1288	AF254411	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
1289	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98
	1	norvegicus		1	

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identit
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine deaminase	468	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis thaliana	ATP-dependent RNA helicase-like protein	636	45
294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22 clone HFEAF41.	504	98
296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	98
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus musculus	Elf-1	806	92
1319	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia coli	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6- PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
1322	AJ276101	Homo sapiens	GPRC5B protein	466	93
1323	AJ276101	Homo sapiens	GPRC5B protein	504	97
1324	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1584	100
1325	U91561	Rattus norvegicus	pyridoxine 5'-phosphate oxidase	1277	89
1326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
1327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2825826.	1531	90
1328	AF151048	Homo sapiens	HSPC214	657	85
329	Y10530	Homo sapiens	olfactory receptor	1645	100
1330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform NaPi-3b	3591	99
1332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
1333	AF078866	Homo sapiens	SURF-4	1395	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	•	Waterman	Identity
NO:				Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

## TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN C7/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino, acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
1	1331	A			_	HWPQAPHRA***GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMRQMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R SCNYTLALILFL
4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  WSHEGEILQAFRGHQGRGIRAIAAHERQAWV ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ
8	1358	Ā	106	3	350	VP**ARYTQGCDSGWLLATAGSD*YRGPVSL *RRGQVLGAAARG*TFPVLLPAGGSSWSRGL RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS WEGAQLELGPAWL FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
					,	LLTEELHILRGVSIYVLRHEAQIYGITPL\VCAL LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY FEEVQRLRFEVHDISSNHNGLKEADFLGGME CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA EELSGNDDYVELAFNARKLDDKDFFSKSDPF LEIFRMNDDATQQLVHRTEVVMNNLSPAWK SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK HDFIGEFTSTFKEMRGAMEGKQVQWECINPK YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI MGGCQIQFTVAIDFTASNGDPRNSCSLHYHHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNPECAGIQGVV EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRG\TLSSSSSSSSLTKDALPSSL KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATS\*G SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPONIVQ MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSOVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNIT\KLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning macleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK GQPHQSTNALLRRCVR*RYHLS\TVETAGLP* KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK
	İ					TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC
				ļ		KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHLLSHIWKDVQNILLK
20	1370	A	304	1	1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA
						LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG
		ł				GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR
	<b>.</b> .	1	ļ	·		LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA
	<u> </u>				}	PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP
				Į		RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP
İ					1	DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP
	l		ļ	! {	Ì	PASPAAASARCPSWGPSCPARPPSRPAAGTEP
	1					AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP   VSFAPEVLSLPAVROTKSWRWRNEEEITRPW
<u>L</u>						ALVRSRGG
21	1371	A	326	799	1587	GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP
			l			HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP
			ł			LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA
			Í		{	GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL
}		ļ	]		ļ	QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA
	<u> </u>			<u></u>	<u> </u>	QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL
			Ì			PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL
		1				SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP
						GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL
						NNEKRKMKKRKEEKKKCRERMQRRSKWRR
		1				EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP
25	1375	A	384	373	128	WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ YLITTILETGYLWKNRHSDQ*KRTENPERDQH
						KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	A	397	383	165	EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI
	<u> </u>	<u> </u>			·	KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFINLTKYQNIQNRHAENYKI
	<u> </u>		<u> </u>			LVNKJEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRP\ANF\*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	A	474	125	471	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEQPASPSPLKPGTYASET\RPRDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR APGGSAGS*\GLPSAGGSRGKKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS* GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRPVVPFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		[	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ì	}	peptide		/=possible nucleotide deletion, \=possible
		L		sequence		nucleotide insertion
						GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
		l	ł		ĺ	GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
						VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
			)	j	j	HGA\GLPGVRVHTQRVHIH*GAG\GCQTPRPR
		l				LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
			]	1		ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
					[	NHQAVGLEASGALQAGHRDELPTMVQLLDH SPDYPLKGRPHAP
20	1207		620	000	<del> </del>	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
37	1387	A	620	828	1	WEASPEMOSKCHOKGKNNQTECFNHVRFLQ
	}		<b>{</b>	١.	ļ	RLNSTHLYACGTHAFOPLCAAIDAEAFTLPTS
						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
	!		į			FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
	]					AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
		1	1	1		SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
		1				\DAEFVFSVLVRESKASAVGDDDKVYYFFTE
		1				RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
"			1.00	-		TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL
	Ì		1	l		KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
	[		ĺ	ĺ	Ì	QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
			1	;		INIIICTVDYLLRLQESI
39	1389	A	767	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
						RNLSVDGKNVDMAGFIANNGTREGCAARRN
}		1	1	ļ		FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
	ļ		1			KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
				1	ļ	VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
	•	1		1		RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
		ĺ	ì	ļ	1	ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
				İ	1	PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
		ļ				LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
]			1	1	i	TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
		ł	ł	i		SEDKVSVRRGFRGCMQVRGGCGGRGEACPS OAAPRL
40	1390	-	801	69	399	UKITIHKEDLNKWKYILCSGMERLSTVMIPVV
40	1390	A	801	09	399	PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL
				1		RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
}	]	ļ	1	ţ	1	VNKSVVLVQVTIM
41	1391	A	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
"'	1371	, A	555	l '	173	SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIIHWIES
72	1372		( 071	1,	1 713	KASFGDECSHHAYLHDQFWSYWNSLKHRTW
1			1	1		QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
1	1	Ì	1	1		FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
!	1	İ	1	{		KACFPTNIVTL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
1 "			1		1 -	LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
	] .		]	1	1	ORPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
	1	1	1			SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
I		1		ſ	-	PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
[	-	1	1	[	1	QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
1		}			1	PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
		ŀ				QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
i .		1	1	1		WLSMSMGK
l	i	1				
46	1396	Ā	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  D/KSLAMLPRLVSNSWPQVILPP  QLQNLASRGCL*SQLLRRLRRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK  HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPTTAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGQYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRŁLESFKEV *LMLA*TKDPSI\RPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLQTRSPSKPLS\RKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

650 TO	CEO ID	NAc4	Tern	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nuci- eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Gondo		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciico	1		***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ	İ	peptide		/=possible nucleotide deletion, \=possible
	1	(		sequence		nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
			1	ŀ		MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
	Ĭ			{	{	TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
		ļ	1	ŀ		ADDTHPARLQGPTLRSQPMGPLKHKAFEERA
		-				NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
	1	1	]			PCLGCPTXATCRLYQTTVAVVF
59	1409	Α	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
		İ	1			KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC
İ	1					SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR
	-	ĺ	Í	(	(	HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
<u></u>		<u> </u>				TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
		1		Ĭ	i	LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
		<u> </u>				EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCOPFHLVMRDLLQLGQDIPQGCHY LEENHLIHRDIAARNCLLSCAAPTRAATIGDF
ŀ			,		İ	
		į .	}		}	GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
		i	l			TN
	1410	<del> </del>	1000	<del>                                     </del>	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER
62	1412	Α	1080	1	839	ANLMHMMKLSIKVLLQSALSLGRSLDADHA
İ	1			1	Į	PLOOFFVVMEHCLKHGLKVKKSFIGQNKSFF
j	1		ì		}	GPLELVEKLCPEASDIATSVRNLPELKTAVGR
	1				ļ	GRAWLYLALMOKKLADYLKVLIDNKHLLSE
		1				FYEPEALMMEEEGMVIVGLLVGLNVLDANL\
			j .			CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
(	1		1			HERITOVLDQKNYVEELNRHLSCTVGDLQTK
l	j	1	1		Į.	IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
1		l	1			REQNELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
03	1 ****	1.	1 .005	] -	""	HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI
		l				HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
			1		ļ	EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
		]	]	ŀ		YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
	1					DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
	1	}		· ·		KAFNHTSICCRHKKN
64	1414	A	1084	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
1		-		ļ		SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS
1	}	1		į.	l	SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
		1		ļ	ļ	LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
	1	1	1	}	1	TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
1		1		1		FFSWLTTGLTTQQRTAIE\NATVAFF\LQCI\SC
		i				HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
Į		1		1		S\GFIR\RLFLQLMLEDEKVTMFLQSPCPLYKG
	1		1		1	RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL
L	<u> </u>			<u> </u>	<u> </u>	SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
						LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA
<u></u>	L				<u>L</u>	SVALHKLSNALV
66	1416	Α	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ
		1		1	Ì	PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
1			1	1	ļ	TLPVAAAFTETVNAYFKGADPSKCIVKITGE
1	1			}		MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
1 -	ŀ	1.				HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL
<u>L</u>				ļ. <u></u>	<u> </u>	MTHLK
67	1417	Α	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
1		1	1	1	1	PYYFLLDLCCSDILRSAICFPFVFNSVKNGST
	1	1	1	1		WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		,	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	<u></u>	<u> </u>		sequence	L	nucleotide insertion
						RYL
68	1418	A	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
	1	1		1		YEREGMQDWKTASGQSEEATQQSSQKPQPH
	ļ	]	i	ļ	ļ	YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
	ì				}	PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
				١.		HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
	ì	1			i	RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
	1	i	ł	į		SPAALAPRAARGGSRAAALAGAEAEEPLRTL
	İ			1		APRPTRAAAPPPPPPPPPPPLPPGAPPPPVRCVSR
		1	ļ	]	ļ	RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
	j	i	}	)	į	APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
			<b>!</b>	1	1	GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
	ŀ	İ				SVIEGVSDQVLVAVVVSFALIATLVYALFRNV HQNIHPENQELVRVLREQLQTEQDAPAATRQ
		ĺ	1		ľ	QFYTDMYCPICLHQASFPVETNCGHLFCGSLT
	1	1	ì			PNSIW
69	1419	A	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
-				] =	100	FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
	)	}	j	j		LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
						TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
		•	1			RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	A	1111	698	23	ALRRLHYVRATKV\FLSFRPFWREEHIEGGH
		}				SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
		ł				AFAGLSREEALRLALDDVAALHGPVVRQLW
						DGTGVVKRWAEDQHSQGGFVVQPPALWQT
						EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
		ł		-		KSALRAAIKINSRKGPASDTASPEGHASDMEG
	1	ĺ	1			QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
	<u> </u>					QNTTHTRTSH
71	1421	A	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
		}	]			PPGPPEQAGLSQFHLEPETQNPETTEEIQSS/LQ
	ļ		ļ			QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
						EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
	[	[				GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI
		ĺ	!			EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
		]	ļ			QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
	}	1				GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
	}					HSYSICHROLKPENLLLDEKNNIRIADFGMAS
		]				LQVGDSLLETSCGSPHYACPEVIKGEKYDGR RADMWSCGVILFALLVGALPFDDDNLRQLLE
			[			KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
	ĺ					LSLEQIQKHPWYLGGNFIS
73	1423	A	1128	1	802	LENALDVLHREVPRVLVNLVDFLNPTIMRQV
				•	302	FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
						MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
						SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
						LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
	[					GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
						TAGLNVAAEGARARDMPAQAWDLVERMKN
						SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
						HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
			_	-		DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
	· '					VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
						EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
		! !				AADPAPVHTTAHPKGA
						I TENDETE VIII TENDERON
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

		134.				T
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		i		peptide		/-possible nucleotide deletion, \-possible
	L		ļ	sequence		nucleotide insertion
						RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ
			1	Ì		DDESGQKKLHGLQAILVHEASGTTAITATAT
İ		1_	i .	i	<u>                                     </u>	GYQESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK
1		ł	1			PDCKEIWIFWWGDEPNLV\VQYIMNCMLWK
1		}	ļ	Į		KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD
			Ì			KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
			1			T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
			1			LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	1	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP
1	- /	ł	}	1 -		SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT
1						TSPRNWCIKMVCNPWFECVSMLVILLNCVTL
	]	1	]	J	]	GMYQPCDDMDCLSDRCKILQVFDDFIFIFFA
1			ļ	ļ		MEMVLKMVALGIFGKKCYLGDTWNRLDFFI
1		1	Ì	1		VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA
ł	ł	ł	ł	ł	ł	INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF
}		1		}		FIFGIIGVOLWAGLERNRCFLEENFTIOGDVAL
j	ļ	ļ	ļ	J		PP\YYQPEEDDEMPFICSLSGDNGIMGCHEIPP
			ì	ļ.	}	LKEOGRECCLSKDDVYDFGAERODLNASGL
ſ		l	İ		1	CVNWNRYYNVCRTGSANPHKGAINFDNIGY
ļ		1				AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI
ł	}	l	Ì	ł	ł	YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP
		1				GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	<u> </u>	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN
19	1429	] ^	1175	<b>,                                    </b>	403	FLSATHLGGLFPPWPLVEERKLKPKASQQCPI
	Ì	l		ţ		CHKVIMGAGKLPRHMRTHTGEKPYMCTICE
		ł	1			VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF
	1	1	ĺ	İ		VKTTRODELENIMEETTOEETTECHENAET VHNYDLKNHMR
	1430	<del>  </del>	1100	25	198	<u> </u>
80	1430	Α	1182	25	198	EMNELSQQLSQQGRGASQCPSPPAPTLPNPT
	1.424	<del> </del>	1106	1054	702	PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS
I		1	1			AIWQQAREVVRFNGLEDRVHVLPGPVETVEL
1		ł	1.	1		PEQVDAIVSEWMGYGLLHESMLSSVLHARTK
	<u> </u>	ļ	·		<u> </u>	VVKDGGFFLPXSSELFM
82	1432	A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA
1 -					1	SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT
	}	ļ	}	ļ	}	SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH
1		1	]	1		GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD
		1		1		TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK
i	1	1		1	1	LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP
1	ł	1	}	1		SQESNVSLSGSSRSALFERDDHGKAEAPSPSF
L	<u></u>	<u></u>	<u> </u>	<u></u>	<u> </u>	DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC
		1	1	!	1	WGRGHGCGQEALSTSHGYHLFCALLTGFLFA
	1			1		SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF
1	1	1		l	<b>₹</b>	Q
84	1434	Α	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES
		1	1	ļ		GACVNQVTVDSITPLHAASLQGQARCVQLLL
]		1	1	1	1	AAGAQVDARNIDGSTPLCECLRLGQHRVCEA
	ļ	1		!	l .	LAVLRGQGQPSPVHSVPPARGLHXREFRMC*
		1		1	[	GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR
		1	1		1	SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ
]	1	ļ	}	1	1	HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT
	}	j	1	1		GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC
55	1.750	1.3		1	.55	NLGFFLEGHAVLTCHAGSENSATWDFPLPSC
1		1	1	1	1	RADDACGGTLRG/AEWHHLQPPLPLG/ATKN
L	1	1	i	1	I	I TO TO SOLD TO THE WILLIAM TO THE POLICIES

NO. of much ceide sed unce where the continual policy of the continual continual policy of the continu	SEQ ID	CEOTO	1 1 6 04	Lego	Dundinted	Deading 1	
nucleotide seq uence view of the peptide celebrate sequence view of the peptide sequence view of the peptide sequence view of peptide residue of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptide sequence view of peptid		SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Seq			поа	1			
sequence  914  914  105  105  106  107  108  108  108  108  108  108  108			l		i .		
1437   A   1216   226   964   37   37   37   37   37   37   37   3		1 -	ļ				
amino acid residue of peptide residue of peptide residue of peptide sequence per per per per per per per per per pe	-	uence				1	
Presidue of   peptide   sequence   y=Tyrosine, X=Ulxhown, *=Siop codes, /possible nucleotide letion,   possible nucleotide insertion   NADCTWILAELEDTIAL YFIDPQLEDGYDFL	uence	[	[	914			
Peptide	ŀ		1	1	amino acid		T=Threonine, V=Valine, W=Tryptophan,
		Į	Į	ļ		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	}	]	peptide		/=possible nucleotide deletion, \=possible
EVTGTEGSLW					sequence	İ	
EVTGTEGSLW				T			NADCTWTILAELGDTIALVFIDFOLEDGYDFI.
1437   A   1216   226   964   GTARFGPMVGFGANRAGRIPSI.VIGUILY		i	ŀ	ł			
	87	1437	A	1216	226	964	GTARFGPMVGFGANRRAGRI PSI VLGVI I V
		1	i	}	}		
		ļ		J	]		
		İ	]	ļ			
			· ·				
EENIKKLADOPLEEQKOETOKIÓSNDOKELDI   NNOVYPKNIPKVAENVADKOELDI   NNOVYPKNIPKVAENVADKOELDI   NNOVYPKNIPKVAENVADKOELDI   NNOVYPKNIPKVAENVADKOELDI   NNOVYPKNIPKVAENTADKOELDI   RVKSAGAWIIHPYSDFRYWDLIMLLIMVGN   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISDIFFLLD   LIVLPVGITFKEENSPPWINFVLISTIFKLIR   LIVLPVGITFKEENSPPWINFVLISTIFKLISTIK   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVLISTIKT   LIVLPVGITFKEENSPPWINFVGITFALLAGAGE   CRIL VERVENTE   LIVLPVGITFKEENSPPWINFWAAN   LIVLPVGITFKEENSPPWINFVGITFKEENSPPWINFVGITFKEENSPPWINFWAAN   LIVLPVGITFKEENSPPWINFVGIT			i	ł			
1438							
1438   A   1218   1   534   PEFGTTISCGYLMATDVSRRPSVEKAVEEGE   RVS.SAGAWIHPTSDERFYUKAVLEEGE   RVS.SAGAWIHPTSDERFYUKALLLANGN   LIVLPVGITFKEENSPHIVPNVLSDTFFLLD   LVLNRTGIVVEEGAELLARARITRYKLTKW   FLVDLISSIPVDYFLVVELEPERLDASVYKTAR   RVDLISSIPVDYFLVVELEPERLDASVYKTAR   ALRIVIPTKLSLILL   ALRIVATIVATE   RVDLISSIPVDADNILLIPDTLIALANTAV   ALRIVATIVATE   RVDLISSIPVDADNILLIPDTLIALANTAV   ALRIVATIVATE   RVDLISSIPVDADNILLIPDTLIALANTAV   ALRIVATIVATE   RVDLISSIPVDADNILLIPDTLIALANTAV   APMLDSRAAYSNFWCGMTSQGYYKETAYI   PIRKRDRRGCFAVPMVHSTFLIDLIALANTAV   APMLDSRAAYSNFWCGMTSQGYYKETAYI   PIRKRDRRGCFAVPMVHSTFLIDLIALANTAV   APMLDSRAAYSNFWCGMTSQGYYKETAYI   PIRKRDRRGCFAVPMVHSTFLIDLIALANTAV   PIRKRDRRGCFAVPMVHSTFLIDLIALANTAV   ANKTSFFYLKNIVADLIMTLTFPRIVHDAGF   GPWDFKFILCKYTSVLFYAMIGUSIVVLGITY   YDRYWKVVRHL/MDSWMTGINSFTRVYLLG   LGARLWFFGKLIAKGGHGHGUISVV.   ALTIVATIVATIVATIVATIVATIVATIVATIVATIVATIVA		]	1	ļ	}	J ·	
RYKSAGAWITHPYSDERFYWDLIMLLIMYGN LIVLPVGITFFKEENSPWINVPNLSDTFFLLD LVLNFRTIGIVVEEGAELLLAPRAIRTRYLRTW FLVDLISSIPYDVIFLVVELEPRLDAEVYKTAR ALRIVEFIKLISLIR.  89 1439 A 1223 1 743 MGFDEVFMINLRRQDDREREMLRALQAGEIE CRIVEAVDGEVGMUTRSNAAPGRILAMLET LVVVAPFFVDADNLILNPDTLSILLIAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCAPFWHYSTFLIDLRKAASRNL VAFYPPHPDYTWSTDDIIVFAFSCKQAEVQMY VCNEEEVGFLPVPLRAHSTLQDEAESFMHVQ LEVMYPSSPSSAQSMAVVSADHIGVISVI.  90 1440 A 1227 2 349 NKTSFFFLKINIVVADLIMTITFPRIVEDAGF GPWDFEKFILCRYTSVLFYANMDTSIVVLGLITI YDRYWKVVRHLWDSWMTGIISFTRVYLLG LGARLWWFGKLILAKGGHGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT GARLWWFGKLILAKGGHGISWL  92 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT GARLWWFGKLILAKGGHGSHAPANLSD VVQLIFLVDSNFFFGYISNYTVSTKVASMAF GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNENNCSASRRIRPESLQGADHRYYTFISFGT RDFVGSYRLNLSSHRRWSALEVSVQLYTSLC QYFSEEDVVRTEGLLIPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPFPTADNYTVML LCAVCL-VTYMMAALHEDOLAASRGRAP FCGGGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHULDGRAFHRNSLDFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGCVEKEVL AASKASSRYVPTSNAALLELEFINGRVSLINPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPLSSAAAR GWGWGPGSTGNGAWASAPCPEPLSS	00	1/20	-	1210	<del></del>	524	
LIVLPYGITFKEENSPPWIVPNVISUIFFLLD	66	1430	^	1210	1	) 34	
September   Sept				1		ļ	
FLVDLISSIPVDVIELVVELEPRLDAEVYKTAR   ALRIVERTIKLISLRL     89		}	1	l	1	Ì	
89 1439 A 1223 1 743 MGPDEVFMINLERRQDEREERMIRALQAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPREVDADNLLINPOTLISLIAENKTVV APMLDSRAAYSINFUNCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASSRIL AFYPPHPDYTWSFDDILVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTI,QDEAESIMHVQ LEVMVPSSPSSAQSMAVVSADHOLVISYU  90 1440 A 1227 2 349 NKTSEIPYLKNIVVADLIMTLTPFFRIVEDAGF GPWDEKFILCRYTSVLFYANMOTSINVLGLIT/ YDRY/WKVVRHLAWDSWMTGI/SFTRVYLG LGARLVWFGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVSQAYNLT SALTPILTRSRVLINEEPHTLAGFSRAPANLSD VVQLIFL/DSNPPFFGYISNYT/STKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAA GHRSSANSVVQPQAFYGAVVTILDSSNPAAV LHLQLNYTLLDGRYLSEEPPYLAVYLHSEPR PNEHNCSASRRIPESCAGDHEPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPSHRRVFPEFTADVNYIVML TCAVCLVTYMVMAAILHKLJOLDASRGRAIP FCGGRGRKYEIL/KTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSIDIFQIATP HSLGSMWKIRVWHDNKGLPAWFLQHIIVPD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLFFRILVAELQRGF FDKHWISJWDRPPSFTRIQAATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVLALLFFRINGRSKC GWGWGPGSTGNGGAWSAPCEPPLSSAAAR GKGYHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDFL REETVYFNCEKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKMYFIQHOVAGVITKSSKEMKLTDGS YFGEICLLYKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVADRLDRIGKKN SILLOKRODLNINGPNNQENELKGUVKH  93 1443 A 1249 180 901 TVPFPFGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSGLPPSSFMVSSAHNPN KAEPPERRKDSTSTFNNLPPSMMTRRNTYYCT		1	1				
89 1439 A 1223 1 743 MGFDEYFMINLRRQORRERMIKALQAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPREVDADNLILNPDTLSLLIAENKTVV APMLDSRAAYSNFWCGMTSGGYYKRTPAYI PIRKRDRRGGCAVPWTSFILDLRKAASRNL VAFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLFVPLRAHSTLQDEAESFMHVQ LEVMYPSSPSAQSMAVVSADHIGLVISVL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFFRIVEDAGF GPWDFEFILCRYTSVLYAMMDTSIVVLGLIT/ YDRY/WKVVRHLWDSWMTGLISTRVYLLG LGARLVWFGKLLAKGGHGGISWL 1441 A 1245 3 1937 LGSSDVRAPQRSELGASPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFFFGYISNYTVSTKVASMAF QTQAGAGPIERLASERPSTWASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFFFGYISNYTVSTKVASMAF QTQAGAGPIERLASERPSTYLVSPNSWARA GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPPTYLAVYLHSPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYNLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTGGLLJEETISPQAVCLTR HLTAFGTSLFVPPSHRRVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLQDLDASRGRAIP FCQQRGRKYELLVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQHATP HSLOSMWKIRVWHDNGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFVYPTSAALLRFERLIVAELQRGF FDKHIWLSIWDRPPRSCFTRQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGUSSVVVVYPLAILFJEFRRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLLGKQOHT  92 1442 A 1246 5 562 VFDEENINENDEPLREGIVFFNCRKLVATMP LFANADPPFVTAM.SKLRFEVFOPGDVIIREG AVGKKMYFQHGVAGVITKSKEMKLTDSG YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVARDLDRIGKKN SILLOKROCDLNITHSVRSCEMKLLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVARDLDRIGKKN SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVARDLDRIGKKN SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS NFNEVLEEYPMRRAFETVARDLDRIGKKN SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLNITHSVRSCEMKLTDGS SILLOKROCDLN		l		1			
CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNLILNPDTLSLIAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNIL AFYPPHPDYTWSFDDIVFAFSCKQWAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISTVI 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVEDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHL-WSWMTGI/SFTRVYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRWASQAYNLT SALTPILTRSRVLNEEPHTLAGFSRAPANLSD VVQLIFL/DSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVK/VPNNSDWAAR GHRSSANSVVQPQAFVGAVVTILDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQADHEPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPFSHRRVFPEFTADVNYIVML TCAVCLVTYMVMAAILHKLQDLDASRGRAIP FCGGRGRRKYEIL/KTGWGRGSGTTAH/GIM LYGVDSRSGHRHLDGDRAFHRNSI DIPQIATP HSLGSMWKIRVWHDNKGLPAWFLQHIVRD LQTARSTIFFLVNDWLSVETEANGGLVEKEVL AASKASFRVFTPSVAALLFFRILVAELQRGF FDKHWISJWDRPPRSTRIQAATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYTLAHLFRMSRSKV GWGWGFGSTGNGAWASAPCEPPLSSAAAR GKGYHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDER EETIVFFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKMYFIQHOVAGVITISSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVADLDRIGKKN SILLOKROPOLNTIGVENELLGVUKH  93 1443 A 1249 180 901 TVPPFPGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGRGLPPSSPMVSSAHNPIN KAEPPERRKDSTSTFNNLPPSMMTRRNTYYCT		1420	<u> </u>	1000			
LVVVAPRFVDADNI.LI.NPDIT.LSLLIAENKT.VV     APMLDSRAAYSNFWCGMTSQGYYKRTPAYI     PIRKRDRRGCFAVPMYHSTFLIDLRKAASRNI.     VAFYPPHPDTTWSFDDIIVPAFSCKQAEVQMY     VCNKEEYGFL.PVPLRAHSTI.QDEASFMHVQ     LEVMVPSSPSSAQSMAVVSADHIGLVISYL     1440	89	1439	A	1223	[ 1	743	
APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL VAYPPPDVTWSTFLIDLRKAASRNL VAYPPPDVTWSTPEDDIIVPAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQAWVSADHIGUISVL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFKFILCRYTSVL,FYANMDTSIVVLGLIT/ YDRY/WKVVRHLUSWMTGIJSFTRVYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESFSRMVASQAYNLT SALTPILITRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNFFFFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGILPLEETSFRQAVCLTR HLTAFGTSLFVPPBHAVDNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHINSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTIFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTISVAALLRFRRLLVAELQRG FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCEPPLSSAAAR GKGVHQRLJGKGWOYPQSTGNACMASAPCEPPLSSAAAR GKGVHQRLJGKGWGAGVTKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVADDRLDRIGKKN SILLQKRQCNDTRISSFWVSSSMVSSAHNPN KAEPPERKDSTSTPNNLPPSMMTRSNTYVCT		1	1				
PIRKRDRRGCFAVPM/HSTFLIDLRKAASRNIL VAFYPPHPDYTWSFDDIDVFAFSCKQAEVQMY VCNKEEYGFLPYPLRAHSTLQDEAESFMHVQ LEVMVPSSPSAQSMAVVSADHIGLVISYU  90 1440 A 1227 2 349 NRTSFPYLKNIVVADUMTLTPFFFYIYBDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHL/WDSWMTGISFTRVYLIG LGARLVWFGKLILAKGGHGGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSVLNEEPLTLAAFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVIVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPFPYLAVYLHSEPP PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHRWSALEVSGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVPFETADNYIVML TCAVCLVTTWMWAALHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHYGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLOSMWKRVWHDNKGLSPAWFLQHILVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTSVAALLRFRRLLVABLQCHGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVYYVYLAILLFRNSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLJGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPL REEIVNFNCRKLVATMP LFANADPNFVTAML SKLRFEVFQPGDYIREG AVGKKMYFIOHQVAVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADDRLDRIGKKN SILLQKRQVIKGSGRGCHPTSSFMVSSSMVSSAHNPN KAEPPERKDSTSTPNNLPPSSMVSSAHNPN		1	l	Į	)	ł	
See		1		ł		}	
90 1440 A 1227 2 349 NKTSPIFYLKNIVVADLIMILTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGISFTRYVLIG IGARLVWFGKLLAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPORSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLILAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTYSTKVASMAF OTOAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVYLIDSSNPAAV LHLQLNYTLLDGRYLSEPFPFYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSQLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNTIVML TCAVCLVTYMVMAAILHKLDQLDASSRGRAIP FCGQRGFKYKEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRINPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWSAPCPEPPLSSAAAR GKCVHQRLLGKOQHT 92 1442 A 1246 5 562 VFDEENILVELNDPLREEIVNFNCRELVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHUVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNQENEELKQIVKH 93 1443 A 1249 180 901 TVPPPPGGFSPAPLHEKRSPTSTGEAELKEERL PGRKASSTAGSSGSTDFNSNMYRSNITPYCT		}			1		
LEMMYPSSPSSAQSMAVVSADHIGLVISYL	l	İ	l	ł	ł	{	\AFYPPHPDYTWSFDDIIVFAFSCKQ\AEVQMY
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVYHL/WDSWMTGISFIRVYLLG LGARL/WFGKLILAKGGHGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANI.SD VVQLIFL/DSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVHQAPVFDAVAVLLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPETTADVNYIVML TCAVCLVTYMVAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFFPTPSVAALRFRRLLVAELQRGF FDKHIWLSWPRPRSCFTRIQRATCCVLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFJERMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNOGENELKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPRRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAINPN KAEDPERKLOSTSTPNNLPPSMMTRRNTYYCT			ĺ			ĺ	VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGISFTRVYLLG LGARLVWFGKLILAKGGHGISWL  91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVBSNPFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPINNSDWAAR GHRSSANSVWQPQABVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLIFLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVPPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCQQRGRKKYELLVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSVVVYPVYLALLFLFRMSRSKV GWGWGGGTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPIPVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTIGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGGAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEDFERKDSTSTPNNLPPSMMTRRNTTYVCT				i	L		
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD VVQLIFLVDSNPTPFGYISNYTVSTKVASMAF\ QTQAGAQIPIERLASERAITVKVPNNSDWAAR\ GHRSSANSV\VQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSARRIRPESLQGADHRPYTFISPGT\ RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC\ QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR\ HLTAFGTSLYPPSHIRFVFPEPTADVNYIVML\ TCAVCLVTYMVMAAILHKLDQLDASRGRAIP\ FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM\ LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP\ HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD\ LQTARSTFFLVNDWLSVETEANGGLVEKEVL\ AASKASFRVPTPSNAALLRFRRLLVAELQRGF\ FDKHIWLSIWDRPPRSCTFTRQRATCCVVLICL\ FLGANAVWYGAVGDSAYSTGRVSRLNPLSV\ DTVAVGLVSSVVYPYYLAILFLFRMSRSKV\ GWGWGPGSTGNGAWASAPCPEPPLSSAAAR\ GKGVHQRLLGKGQHT\  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP\ LFANADPNYTAMLSKLRFEVFQPGDYIIREG\ AVGKKMYFIQHGVAGVTIKSSKEMKLTDGS\ YFGEICLLTKGRRTASVRADTYCRLYSLSVD\ NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN\ SILLQKFQKDLNTIGVFNQENELKQIVKH\  93 1443 A 1249 180 901 TVPPPGGPSPAPLHPKRSFTSTGEAELKEERL\ PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN\ KAEPPERKKDSTSTPNNLPPSMMTRRNTTYCT	90	1440	A	1227	2	349	
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD VVQLIFLVDSNPTPFGYISNYTVSTKVASMAF\ QTQAGAQIPIERLASERAITVKVPNNSDWAAR\ GHRSSANSV\VQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSARRIRPESLQGADHRPYTFISPGT\ RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC\ QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR\ HLTAFGTSLYPPSHIRFVFPEPTADVNYIVML\ TCAVCLVTYMVMAAILHKLDQLDASRGRAIP\ FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM\ LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP\ HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD\ LQTARSTFFLVNDWLSVETEANGGLVEKEVL\ AASKASFRVPTPSNAALLRFRRLLVAELQRGF\ FDKHIWLSIWDRPPRSCTFTRQRATCCVVLICL\ FLGANAVWYGAVGDSAYSTGRVSRLNPLSV\ DTVAVGLVSSVVYPYYLAILFLFRMSRSKV\ GWGWGPGSTGNGAWASAPCPEPPLSSAAAR\ GKGVHQRLLGKGQHT\  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP\ LFANADPNYTAMLSKLRFEVFQPGDYIIREG\ AVGKKMYFIQHGVAGVTIKSSKEMKLTDGS\ YFGEICLLTKGRRTASVRADTYCRLYSLSVD\ NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN\ SILLQKFQKDLNTIGVFNQENELKQIVKH\  93 1443 A 1249 180 901 TVPPPGGPSPAPLHPKRSFTSTGEAELKEERL\ PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN\ KAEPPERKKDSTSTPNNLPPSMMTRRNTTYCT			ļ				GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAWVGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGFSPAPLHPKRSPTSTGEAELKEERL PGRKASCTAGSGSRGLPVSSPMVSSAHNPN KAEIPERKDSTSTPNNLPPSMMTRRNTYVCT	İ						
SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPGYISNYTYSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIPESLQGADHRPYTFISFGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDFLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGFSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGRGLPPVSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSSMMTRRNTYVCT		1		1		İ	LGARLVWFGKLILAKGGHGGISWL
SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFFPTADVNYIVML TCAVCLVTMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTIFFLVNIDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERKDSTSTFNNLPPSMMTRRNTYVCT	91	1441	A	1245	3	1937	LGSSDVRAPQRSELGAESPSRMVASQAYNLT
VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLIPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTSNAALLFRRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPSTGINGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		Į				į	SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD
QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFYGAVVILDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRFESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASSGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERKKDSTSTPNNLPPSMMTRRNTYYCT			1				VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF
GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEFPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEFTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFFVFQPGDVIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERKKDSTSTPNNLPPSMMTRRNTYYCT		l		}			OTOAGAOIPIERLASERAITVKVPNNSDWAAR
LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHENKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLALFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPFGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVYCT		İ		-		ł	
PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGA WASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPALHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	}	ł	i	ł		i	LHLOLNYTLLDGRYLSEEPEPYLAVYLHSEPR
RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYYCT			ľ				
QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGGGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYYCT		1	ļ				RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPVSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		]	į				
TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGGRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1					
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			l				
LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	1	l			l	
HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		ł	1		1		
LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGA WASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		]	1	1 .			
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1				
FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT  92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1				FOR HIM CHAD DDD COLLEGE ATOMS COL
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1			}	FI CANAVAYOR AND AND TO THE CANAVAYOR AND THE CANAVAYOR AND CANAVAYOR AND CANAVARDA AN
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		l	l				
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			Į.				
LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	-	L	<del> </del>	10/5			
AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH  93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	92	1442	A	1246	3	<b>&gt;62</b>	
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	l				
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	1	l	}		l	
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	i				
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
93   1443   A   1249   180   901   TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL   PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN   KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		<u></u>	L		<u></u>		SILLQKFQKDLNTGVFNNQENEILKQIVKH
PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	93	1443	Α	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1	<b>]</b>			
			1	] .			KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT
		1	1				ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
	1444		1001	3	205	TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP RQEDHLSPGGRGCSEL
94	1444	A	1261	3	385	KFSQWGLTKPKLSNASP/WISLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR
95	1445	A	1282	2	550	GPRDNPGVEDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILKLARH STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA
96	1446	A	1294		1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKRKQKEQSGEEKDED EFQKSESEDSIRRKGFFFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97	1447	A	1295	2	2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HIKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
98	1448	A	1304	118	453	SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
99	1449	A	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	aciioc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice			214	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			}	peptide	sequence	/=possible nucleotide deletion, \=possible
			·	sequence		nucleotide insertion
<b>—</b>		<u> </u>		Sequence		DSRLLLFDAPDLRLSPPSGALLQVLDLRDPOF
i		1	1			
		[	[			SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP
	ļ				]	TTCTVLLLAESEGERERWLQVLGELQRLLLD
						ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD
		ļ				QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ
-						QLTLSPSAGLLVVLCGRGPSVRLFALAELENI
			i .			EVEVPKIPESRGCQVLAAGSILQARTPVLCVA
		1				VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
		ł	1	1		SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL
		İ				GAGLVPEELPPSRGGLGEALGAVELSLSEFLL
i			1	Į	1	LFTTAGIYVDGAGRKSRGHELLWPAAPMGW
		1				GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL
1				ŀ		KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE
}		<b>}</b>		1	,	KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE
						EQQKQQRREMLKDPFVRSKLISPPTNFNHLV
100	1450		1310		100	HVGPANGRPGARDKSP
100	1450	Α	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA
1		ļ	}	-		LNDSMINETARDAARVQVASTLSVLVGLFQV
1						GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF
			Ì			VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL
<u> </u>		Ì	ļ	ļ	·	PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ
1		[				QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\
ļ		•	J .			PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK
101		ļ. <u>,                                    </u>	10.55			IFALRHGYRVDSNQVWVMRDV
101	1451	Α	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR
1		1		į		SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD
100	1.00	<u> </u>	10.00			HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN
ĺ		i				WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY
						CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA
1		l	(	ĺ		VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW
		ł				YNLILTA/LLCLPLVIVTLCYTTIIHTLTHGHAN
100	1150		10.54			\DSCLKQKARRLTILLL
103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
				ł		SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
						GYQLYDVCSDSANVYATLRVLSLPGQHHIEL
1			İ	ļ		QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP
1.04						FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
		1				FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP
		}	Į	Į		RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW
		[		Į		ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG
100		<del> </del>		<u> </u>		KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD
		ŀ	1	1		IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF
	ļ	İ	1			SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK
		[	[	[		CQSDWMERWVDDAFWSFLF\SLILIVIMFLW
					L	RPSA
106	1456	Α	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW
		[	1			QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG
		İ		l		CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL
		ł	{	ł		FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF
	<u> </u>	<u></u>	<u>L</u>	<u> </u>		QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN
		L_				YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM
1	ŀ	l	ļ			VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT
	1	ł	1			VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK
	L	<u>L</u>		<u> </u>		AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG FTYNVGALGGALAPIIGALIAQRLDLGTALAS LSFSLTFVVILRNRRPGKSLVR
109	1459	A	1402	15	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS WSHNSNSMCWGKDQCPYSGCKEALIRTDGM RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS GVYCCRIEVPGWFNDVKINVRLNLQRASTT
110	1460	Α	1421	3	350	HEDLSSLLTRGSGNQERERQLKKLISLRDWM LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST SDSLLFSQDSKLATTS
112	1462	A	1434	46	372	TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC MSSSTTSSTTSTF
113	1463	A	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLIVCVLLFYVSFYLFQSCINFVL
114	1464	A	1463	1	396	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS EDP*KNA*LKQMHAATTHWQQHQQHQVGC QYHGIMQ
115	1465	A	1464	291	2	AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN
116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC
117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK
118	1468	A	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H
119	1469	A	1486	1	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL
120	1470	A	1497	3	999	MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

PCT/US01/03800

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Amino add a second
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ	Ì	peptide		/=possible nucleotide deletion, \=possible
			ŀ	sequence		nucleotide insertion
						MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
	·	1				GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
						WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE
		İ				MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD
_			]			SSDFVEFEN
122	1472	Α	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT
	j	l	1			WDPGTDTALGWSKQPSQSYTLFES*VGSGYII
			L		ļ	DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
ŀ		<u> </u>				RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA
						AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
						HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
l l		1			j	HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP
						YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
İ						HLYEERGGVIDITAWDKDAGKRDDFIGRCQV
		ľ				DLSALSREQTHKLELQLEEGEGHLVLLVTLT
		1				ASATVSISDLSVNSLEDQKEREEILKRYSPLRI
		ł				FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
						PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
		İ				*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
						ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY
J		}				KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
						AWDKDAGKRDDFIGRCQVDLSALSREQTHK
						LELQLEEGEGHLVLLVTLTASATVSISDLSVN
						SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV
						KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
125	1475	A	1556	57	509	THTVYKNLNPEWNKVFTL
125	1473	Λ.	1550	37	309	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
						CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT
						TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
						KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	A	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL
	,.	••	10,2	3	170	EMLPTCDLADQHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
				-		VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
						FWPETEKPKITLKNAMKMESGDSGNLL*AAT
1				1		QGASSSISLVANIAVNLIAFLALLSFMNSALA
						WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
						WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA
1						EDEVDFRASSISEEVAVGSIAATLKMKQGPM
						TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
j			j			MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA
ľ						LLCVWALSLVTYIGPLLGWRHPAPEDETICOI
j						NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
			[		<b>f</b>	AKTE
130	1480	Ä	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG
İ	1	i	l			ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
ţ	'					EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP
1		,	ł			KPSEASTTLPESHLKTPOMPPTTPSSSSFTKVT
						KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE
131	1481	A	1651	607	3	AIRYCHGCGVAHRDLKCENALLOGFNLKLTD
131	1481	A	1651	607		LAEGGDVFDCVLNGGPLPESRAKALFRQMVE AIRYCHGCGVAHRDLKCENALLQGFNLKLTD FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TDIPKMLWQQQKGVSFPTHLSISADCQDLLK RLLEPDMILRPSIEEVSWHPWLAST**KQWQV LSNKVGGESKPKKKK
			_			EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV VDAQ
133	1483	Α	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	Α	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSSPRDRDRER*RTRERERERDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIOYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	A	1678	525	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
137	1487	A	1680		2999	AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSQKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHPSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST PEEFSRRLKERMHHNIPHRFNVGLNMKATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRGQQGWDRK YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

ero m	SEC ID	Met	SEQ	Predicted	Dradiated	Amino opid populare (Am Alania Ca Carati
SEQ ID NO: of	SEQ ID NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иелсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
	_	}		sequence	ĺ	nucleotide insertion
						YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
		ļ				NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
		j			}	ISSGAIYLASSYQDKLRVICCKGNLVKESGTE
				<u> </u>		HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	A	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
		•			{	PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
		}	j		j	PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
			1			CWTRGCQTTARTAAAAAAPGPAGRRPPGGA
				ļ		PQNGSCAASASQEAAAPPPMCPPGRRWAVAS
139	1489		1693	3	276	PPETRCPAAPGTRCRRLEAA
139	1469	A	1093	٥	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
			Ì			FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
				į		IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
140	1490	A	1704	3	376	RWAGIAKGVGTQKIIGRVHLGEQKALGL ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA
140	1,470	<b>^</b>	1707	1	1 370	HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
					ļ	KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
						LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
				_	1 302	DKLELELVLKGSYEDTQTSFLGTASAFRFHY
			Ĭ		i	MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
		l	ĺ	ĺ	(	PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	1	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
		•			[	LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
					ļ	GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
			Ì		,	LLQVGDRVLSINGIATEDGTMEEANQLLRDA
			<u> </u>	<u></u>		ALAHKVV
143	1493	A	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
						KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
						NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
		ł				SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
144	1404	<u> </u>	1014		10.1	LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
		ļ				PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
		ĺ	<b>(</b>			KCNGEWVSQNDHVTQEGLDEATGLRVREVH
	İ	]				IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
		1				CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
						THLALCPIVQHPEDTCIHSREVGVVCSRYTDV
		ł				RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
						PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
						SMAAET*HHVPASGADPYVRVYLLPERKWA
		l	}			CRKKTSVKRKTLEPLFDET
147	1497	A	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
		l	}			VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
		1				TSVTYSMG*HGAPTGSEAGANWNH**LHAH
		l <sup>.</sup>				YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
						Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
		l	]			IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
	Ì	1				GIEGRLTADQLNSATACIFAAEVAIKESERFN
		l			]	GIPALSVPVAEPIRHAEALMQQALTLKRSDET
		l				RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
146	1455	<del> </del>	1000			VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
						PSQIRVVATATLRLAVNAGDFIAKAQEILGCP VQVISGEEEARLIYQGVAHTTGGADQRLVVD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	, nou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	Jones	]	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
dence		1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ŀ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		}		peptide	acduence	/=possible nucleotide deletion, \=possible
				,		
	<u> </u>	<u> </u>	<u> </u>	sequence		nucleotide insertion
						IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
			Ì			YFADRNLGQENFDAAQKAAREVLRPVADEL PAULON CONTROL OF THE CONTROL OF TH
}						RYHSWKEVRGASVTVQALQEIMMAQGMDE
						RITMEIWPVD
150	1500	Α	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
				'		LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
[						EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
						PESSILDGMIRQLQQQQDQRMGADQDTIPRG
			l i			LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV
		ĺ	[			EGVRQMHQNAPRSQIATERDLQAWKRRVVV
						PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
			<u></u> _		<u></u>	TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA
						MEMQIKKQFQDTCKVQTKQYKALKNHQLEV
}						TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
				'		NEMMASQALRLDEAQEAECQALRLQLQQEM
						ELLNAYQSKIKMQTEAQHERELQKLEQRVSL
1 1						RRAHLEQKIEEELAALQKERSERIKNLLERQE
			i '			REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	Α	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL
						RQKIFKERALPDIENYMFENHDQLRQAATEC
						MCNMVLHKEVQERFLADGNDRLKLVVLLCG
						EDDDKVQNAAAGALAMLTAAHKKLCLKMT
1						QVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL
				-		ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG
]			'		]	FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
]				~	1	YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
	.					RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
					'	QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA
100	כטכו	Λ.	1747		307	KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
						LLLKDMDLLRGREAELKQRVEAFESYQLELK
156	1506	· 🛪	1025		270	DDYIRTYRLIEDDRINIQISGHWQESP
156	1506	A	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADČLLENEL
						DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
157	1500	<u> </u>	1036	504	705	RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
	'		·			AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
			10-0			NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA
					ľ	NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
						RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
						QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
						RFSRFAGSSPIQSSM
159	1509	Α	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
						NGHLATVKLLVEEKADVLARGPLNQTALHL
						AAAHGHSEVVEELVSADVIDLFDEQGLSALH
				,	·	LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
1						HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
{ }		,			-	IYSEYCNNHPGACLELANLMKQGKYRHFFEA
						CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL
						LKYTTOEHGDYSNIKAAYEAMKNVACLINER
1						KRKLESIDKIA
161	1511	Ā	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
'''	7214	FL.	1707	*	.,,	LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
						SSHNASKTLNANNMETLIECQSEGDIKEHPLL
L					L	COLUMN TENAMINITURE LE LECCORDINATION DE

nucleotide sequence unne	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:			D=Aspartic Acid, E=Glutamic Acid,
1949    1949				1			
1914   ng to first amino acid residue of peptide pep							
minto acid residue of peptide residue of peptide sequence   T=Therentiae, N=Valine, N=Typophan, V=Typosito, Sequence   Sequence   T=Therentiae, N=Juliae,		uence	ļ	1			
Persidue of   Peptide   Sequence	uence	ł	ł	914			Q=Glutamine, R=Arginine, S=Serine,
peptide   sequence	1		ł				
	1	j	1			sequence	
ASCESEDSICQLEWKRKKVLSWFFIARRES   PASSPSGALETDIKASLPOPLSINGESSDTLP   PASSPSGALETDIKASLPOPLSINGESSDTLP   PASSPSGALETDIKASLPOPLSINGESSDTLP   PRIQUILTILCLKGPSTEGIFRAANEKARKEL   SEEINSGDAVDLERLYPHLAVVENDETERS   RKLLSSDLFEEWMGALEMQDEEDREALK   LENGGLFSAPPSSANEEMELTIGGONWCSGIEL   KFQGWGTYCODMFNIDHASVICRQLEGGSA   VSFSGSSNFGEGSGPWFDDLICKGNESALWN   CKHQGWGTYCDDMFNIDHASVICRQLEGGSA   VSFSGSSNFGEGSGPWFDDLICKGNESALWN   CKHQGWGKHQCDIABADAVICSSSLE   KFQGWGTYCDDMFNIDHASVICRQLEGGSA   VSFSGSSNFGEGSGPWFDDLICKGNESALWN   CKHQGWGKHQCDIABADAVICSSSLE   KFQGWGTYCDDMFNIDHASVICRQLEGGSA   VSFSGSSNFGEGSGPWFDDLICKGNESALWN   CKHQGWGKHQCDIABADAVICSSSLE   KFQGWGTYCDDMFNIDHASVICRQLEGGSA   VSFSGSSNFGEGSGPWFDLICKGNESALWN   CKHQGWGKHQCDIABADAVICSSSLA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA   CKHQGWGKHQCDIAGANATODIA		1		· '			
PASDFSGALETDLKASLFDOPLSIIGGDSDTLP   RPIQDLTILCLKGFSTEGIFRAANKEL   KEELNSGDAVDLERLPYHLLAVYKDFLRSIP   RPIQDLTILCLKGFSTEGIFRAANKEL   KEELNSGDAVDLERLPYHLLAVYKDFLRSIP   RRLISSDFEBWAGLERLPYHLLAVYKDFLRSIP   RRLISSDFEBWAGLERGOPMERALK   KEELNSGDAVDLERLPYHLGAVYKDFLRSIP   RRLISSDFEBWAGLERGOPMERALK   KEELNSGDAVDLERLPYHDLAVYKDFLCSSA   VSFSGSSNFGEGSGPWFDDLLCNGNESGRUF   KPQGRWGTVDDDHYDDHASVICHQLCCSA   VSFSGSSNFGEGSGPWFDDLLCNGNESALWN   CKHQCWGTVDDDHYDDHASVICHQLCCSA   VSFSGSSNFGEGSGPWFDDLLCNGNESALWN   CKHQCWGTVDDDHYDTSKYTAPQAATNGD   LASKSNIAFMGTLVRCGRAKGVVIGTGENSE   FGDINLSTTVVHS   FGDINLSTTVVHS   SLLCLFFGTSTVVCKPIVETQLYVIVAQLFGG   SHYKKDSFANKFIKQAIEILKRKRNDIETTKI   ENNWYTVVADSKAGFTTTYKWERCTGYSH   SHYKKDSFANKFIKQAIEILKRKRNDIETTKI   ENNWYTVVADSKAGFTTTYKWERCTGYSH   NOTHWEVAHERDELWRAQIVATTVMLERKLP   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGDRWLRVEDRQDLN   RORIQRYA   RCLWPRSGICGREFGGRWLRVEDRQDLN   RORIGRAM   RCLWPRSGICGREFGGRWLRVEDRGFGGSSKVFLTY   SYLLAPIVULLPYFIPASALFFRVGFVVAER   VLYBFSMGYCLPYPPRASALFFRVGFVVAER   VLYBFSMGYCLPYPPRASALFFRVGFVVAER   VLYBFSMGYCLPYPPRASALFFRVGFVVAER   VLYBFSMGYCLPYPPRASALFFRVGFVVAER   VLYBFSMGYCLPYPPRASALFFRVGFVVAER   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGICGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGREFGGADAV   RCLWPRSGIGGR		ļ	<del> </del>	<del> </del>	sequence		
RPIQDILITI.CLKGPSTEGIFRANDKARKEL   REELINGDAVDLERIPYILLAVYKDFERSIP   RKLLSSIDLFEWMGALEMQDEEDRIFALK   REELINGDAVDLERIPYILLAVYKDFERSIP   RKLLSSIDLFEWMGALEMQDEEDRIFALK   RKPQGRWGTVCDDNRNDHASVICRQLEGGSA   VSFGSGSNRGEGSGPIWPDDLICRGNMCSGREI   KPQGRWGTVCDDNRNDHASVICRQLEGGSA   VSFGSGSNRGEGSGPIWPDDLICRGNMCSGREI   KPQGRWGTVCDDNRNDHASVICRQLEGGSA   VSFGSGSNRGEGSGPIWPDDLICRGNESALWN   CKHQGWGKINCDHAEDAGVICSSKD   CKHQGWGKINCDHAEDAGVICSSKD   AVDLSIDESSIT GETTFCSKYVTARPQRATNGD   LASRSNIARMGTILVECCKAKGVVIGTGENSE   FODIDILSTFVVIRS    FODIDILSTFVVIRS    LASRSNIARMGTILVECCKAKGVVIGTGENSE   FODIDILSTFVVIRS    FODIDILSTFVVIRS    FODIDILSTFVVIRS    FODIDILSTFVVIRS    FODIDILSTFVVIRS    LASRSNIARMGTILVECCKAKGVVIGTGENSE   FODIDILSTFVVIRS    FODIDILSTFVIRS    FODIDILSTFV		1		!			
	ļ				İ		
RKLLSSILFEEWMGALEMQDEEDRIEALK	ì	ľ		ľ			
1512   A   1986   864   501	ļ	}	ŀ				
	162	1512	<u> </u>	1086	264	501	
VSFSGSSNFGEGSCPWFDDLICNGNESALWN   CKRIGGWGKNENDCHAEDAGVISSKD	102	1312	1	1700	804	301	
163	j	· ·	]				
1513   A   2001   419   187	1	ļ	1				
LASRSNIAFMGTI / PRCGKAKGVVIGTGENSE   FGDINLSTEVYHS	163	1513		2001	419	197	
164	103	1213	^	2001	419	107	
1514	1						ECDINI CTENNIC
SHIYKRDSFANKFIKIQAELIKIRKPRDÜETTKI	164	1514	Δ	2012	284	507	
ENNWYPVADSSKAGFTTIYKWERETGFYSH QSFTR	104	1314	1 ^	2012	204	J71	
165		<b>l</b> .	ŀ				
1515	1				}		,
NVDLPFMYSITY-AAFAITLMINLIAMMG   DTHWRVAHERDEL WRAQIVATIVMLERKLP   RCL.WPRSGIGGREYGLDRWILRVERDQLIN   RQRIQRYA   RQRIQRYA   CCQREGIGIKAVVQILLSHGRNGLPGEPASS   QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF   NKNMVTERLQNVMVLEQCFSDSSSL YRFLTY   SYLLAFNVWLLAFVTLCYDWQVGSJELVETI   WDMRNLATIFLAVVMALLSLHCLAAFKRLE   HKEVLVGLIFLVFYPFSNLFRVGGVVAER   VLYMPSMGYCILFVHGLSKLCTWLNRCGATT   LIVSTVLLLLIFSWKTVKQNEIWLSRESLFRS   GVQTLPHMAKVHTYNYANFLKDQGRNKEAIY   HYRTALNNNKAWDYLCWRFRKTLTDLP   AASAASSLTVTTGALSACSHSILRSGPGA   ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW   PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY   GRIFAVVHFASRQWKYTSEDLLIGNELDLA   CGERRLEKVLLVGADNFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKKRKVKKIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADNFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKKRKNFKKKRIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADVFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKRKNFKKKRIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADVFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKRKNFKKKRIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADVFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKRKNFKKKRIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADVFTLLGKPLLGKDLV   RVEATVUEKTSSWPHINFKRKNFKKKRIV   TTPQTVLRINSIEIAPCLL   ACGRIRALEKVLLVGADVFTLLGKGQAQAPQ   RLQGAARVFMPLQAQVKARASKPLQAQAPQ   RLQGAARVFMPLQAQVKARASKPLQAQAPQ   RLQGAARVFMPLQAQVKARASKPLQAQAPQ   RLQGAARVFMPLQAQVKARASKPLQAQAPQ   RLQGAARVFMPLQAQVKARASKPLQAQAPQ   RUQGAARVFMPLQAQVKARASKPLQAQAV   RUGGARGANGASKPLVPRLGGGKPS   SPSLSFLMGFGSNKNAKKKINDUVKTFA   QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS   VIGSLNIFONDOVLKTKVPPVHGKSGLKT   NLTGTDSPETDASSASLQQNKSPRVPRLGGKPS   SPSLSFLMGFGSNKNAKKKINDUVKTFA   QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS   VIGSLNIFONDOVLKTKVPPVHGKSGLKT   RUTTUESGEGGPLGFFLDRINGRGSH   RUTTUESGEGGPLGFFLDRINGRGSH   RUTTUESGEGGPLGFFLDRINGRGSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGPLGGAARASKECVLNE   RUTTUESGEGGPLGGHPLFRELVMFRGSSH   RUTTUESGEGGFLGGGAARASKECVLNE   RUTTUESGEGGFLGGGAARASKECVLNE   RUTTUESGEGGFLGGGAARASK	145	1515	_	2012		402	
DTHWRVAHERDELWRAQIVATTVMLERKLP   RCLWPRSGICGREYGLGDRWILRVEDRQDLN   RQRIQRYA	103	1313	^	2013	2	403	
RCLWPRSGICGREYGLGDRWILRVEDRQDLN RQRIQRYA  166 1516 A 2019 2 927 CCQREGLGLKAVVQILLSHGRNGLPGEPASS QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF NKMWTERLQNVMVLEQCFSDSSLYRFLTY SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI WDMRNLATIFLAVVMALLSHCLAAFKRIE HKEVLVGLIFLVFPFSNLFFRVGFVVAER VLYMPSMGYCILFVHGLSKLCTWLNRCGATT LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNANRLRDQGRNKEAIY HYRTALINNIKAWDYLCWRFRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGFGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFA-VVHFASRQWKVTSEDILLIGNELDLA CGERRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRINFKKKRIV TTPGTVLRNISEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVASPRLLQVGSQVS KKQQAQTOTSEPQDLDQVPEFFGQDQQVLR PPLRRAARVLMPLQAQVRAPRLLQVGSQVS KKQQAQTOTSEPQDLDQVPEFFGGQDQVLR AQDVFRQAMKSPSVLLHVLPPYNFSSLSGRILGFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NILTGTDSFETDASASLQQNKSFRVPRLGGKPS SPSLSPLMGFGSNKAKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTOEELVAMRSTRQG ETASLVARQEGHFPELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDK TLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAPVLPGLSPLIVEEDAV ESSSESSLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDFAGPLFROKEPIODFLDPS DAQFVDVHISDTIAALGYKEPLGNDFYPNGG	İ			İ	'		
RORIGRYA	1						
1516	İ		•		!		
QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF NRNMVTERLQNVMVLEQCFSDSSLYRFLTY SYLLAPNVWLLLAPVTLCYDWQVGSIPLVETI WDMRNLATIFLAVVMALLSLHCLAAFKRLE HKEVLVGLLELVFFIPSANLFFRVGFVVAER VLYMPSMGYCILFVHGLSKLCTWLNRGGATT LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYAPILKDQRNKEAIY HYRTALNNNKAWDYLCWRPRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSVLFGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIERTESPPRIIMFRRKNFKKKRIV TTPQTVLRINSIELAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMGIKA PPRLRRAARVLMPLQAQVKAKASKPLQMGIKA PRUBEVLINGVQTELLTSPRTGDTLSSMTR TVEISGEGGPLGHIVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPTDASASLQONKSPRVPRLGKPS SPSLSPLMGFGSNKNAKKIKDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PPVATHLTKLINSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVADAVASKCVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITDLDFAGPLFNGKPHQDRLDPS DAQFVDVINISDTDALGYKEPLGNIDFYPNGG	166	1516		2010	2	027	
NKNMYTERLQNYMVLEQCFSDSSSLYRFLTY   SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI   WDMRNLATIFLAVVMALLSLHCLAFKRLE     HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER     VLYMPSMGYCILFVHGLSKLCTWLNRCGATT     LIVSTVLLLLLLESWKTVKQNEIWLSRESLFRS     GVQTLPHNAKVHYNYANFLKDQGRNKEAIY     HYRTALNNNKAWDYLCWFFKTLTDLP     1517	100	1510	1 ^	2017	2	921	
SYLLAPNWLLLAPVTLCYDWQVGSIPLVETT   WDMRNLATIFLAVVMALLSHCLAAFKRLE   HKEVLVGLIFLVFPFIPASNLFFRVGFVVAPR   VLYMPSMGYCILFVHGLSKLCTWLNRCGATT   LIVSTVLLLLLESWKTVKQNEIWLSRESLFRS   GVQTLPHNAKVHYNYANFLKDQGRNKEAIY   HYRTALNNKAWDYLCWRFRKTLTDLP     167   1517   A   2025   696   71   AAASAASSLTVTLGRLASACSHSILRPSGPGA   ASLWASRRFNSQSTSYLPGVYPKTSLSSPPW   PEVVLPDPVEETRHAEVVKKVNEMIVTGQY   GRLFAVVHFASRQWKYTSEDLLLIGNELDLA   CGERIRLEKVLUVGADNFTLLGKPLLGKDLV   RVEATVIEKTESWPRIIMRFRKRNFKKRIV   TTPQTVLRINSIEIAPCLL     168   1518   A   2046   2   366   HLQVAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGQAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAPQ   RLQGAARVFMPQAVDSAPKPLKGGAQAQAPQ   RLQGAARVFMPLQAQVKAKASKPLQMCKA   PPRLRRAARVLMPLQAQVKAKASKPLQMCKA   PPRLRRAARVLMPLQAQVKAKASKPLLDVPQGQVLR   RUGSLNIRGNNGQTLKTNPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPVHGKSGLKTA   RUGSLNIRGNNGQVLKTKVPPVHGKSGLKTA   RUGSLNIRGNNGQVTQVLTRSSQKA   RUGSLNIRGNNGQVLKTRSSAKTLCETVKDF   VAKVEKTYDKTLENAVVADAVASKCSVLNE   KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV   ESSSEESLGSKRQLGDDVTKPSSQKA   PSRPWLGRITGLDPAGPLFNGKPHQRRLDPS   DAQFVDVHINSTDALGYKFPLGNDFYPNGG   DAQFVDVHINSTDALGYKFPLGNDFYPNGG   DAQFVDVHINSTDALGYKFPLGNDFYPNGG			ł		·		
WDMRNLATIFLAVVMALLSLHCLAAFKRLE HKEVLVGLLFLVFPIPASNLFFRVGFVVAER VLYMPSMGYCILFVHGLSKLCTWLNRCGATT LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYANIFLKDQGRNKEANY HYRTALNNKAWDVILCWRFRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPPDYETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLLIGNELDLA CGERRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMFFRKRNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVVSAKKRIVCHQQQQAPQ RLQGAARVFMPLQAVVAKAKSKPLQMQIKA PPRLRRAARVLMPLQAQVYAPRLLQVGSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGLGHVVPFFSSLSGRILGFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSSRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSHGPGFIVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKGDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHISDTDALGYKEPLGNIDFYPNGG	}		1	}			
HKEVL/GLLFL/PPFIPASNLFFR/GFVVAER VLYMPSMGYCILFVHGLSKLCTWLNRCGATT LIVSTVLLLLIFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYANFLKDQGRNKEAIY HYRTALNNNKAWDYLCWRFRKTLTDLP 167 1517 A 2025 696 71 AASAASSITVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRILFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRNFKKKRIV TTPOTVLRINSIEIAPCLL 168 1518 A 2046 2 366 HLQVAARVFMPLQAQVSAKASKPLQMQIKA PPRLRAARVHMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQQDQVLR 169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRLGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ GGGRILEVNGRDVTGRTOEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKLINSDEHAVVISSAKTLCETYKDF VAKVEKTYDKTLENAVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESIGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHISDTDALGYKEPLGNIDFYPNGG	1			} ,			
VLYMPSMGYCILFVHGLSKLCTWLNRCGATT   LIVSTVLLLLIFSWKTVKQNEIWLSRESLFRS   GYQTI-PHNAKVHYNYANFLKDQGRNKEAIY   HYRTALNNNKAWDYLCWRFRKTLTDLP     167   1517   A   2025   696   71   AAASAASSLTVTLGRLASACSHSILRPSGPGA   ASLWSASRFNSQSTSYLPGYVPKTSLSSPPW   PEVVLPDYVEETRHHAEVVKKVNEMIVTGQY   GRLFAVVHFASRQWKVTSEDLILIGNELDLA   CGERIRLEKVLLVGADNFTLLGKPLLGKDLV   RVEATVIEKTESWPRIIMRFRKRNFKKKRIV   TTPQTVLRINSIEIAPCLL     168   1518   A   2046   2   366   HLQVAARVFMPLQAQVKAKASKPLQMQIKA   PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA   PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA   PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS   KKQQAQTQTSEPQDLDQVPEEFGQGDQVLR   CONSRSKREGLFHENECIVKINNVDLVDKTFA   QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS   VIGSLNIFGNDGVLKTKVPPPVHGKSGLKTA   NLTGTDSPETDASASLQQNKSPVPRLGGKPS   SPSLSPLMGFGSNKNAKKKIDLKKGPEGLGF   TVVTRDSSHGPGPIFVKNILPKGAAIKDGRLQ   SGDRILEVNGRDVTGRTQEELVAMLRSTKQG   ETASLVIARQEGHFLPRELVMFRSQSH   ETASLVIARQEGHFLPRELVMFRSQSH   ETASLVIARQEGHFLPRELVMFRSQSH   ETASLVIARQEGHFLPRELVMFRSQSH   ETASLVIARQEGHFLPRELVMFRSQSKA   ETASLVIARQEGHFLPRELVMFRSQKA   ESSESESLGESKEQLGDDVTKPSSQKA   170   1520   A   2050   363   1   PVATHLTKILNSDEHAVVISSAKTLCETVKDF   VAKVEKTYDKTLESNAVVADAVASKCSVLNE   KLEQLLQALHTDSQAAPVLPGLSPLIVEDDAV   ESSESESLGESKEQLGDDVTKPSSQKA   171   1521   A   2055   139   675   IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS   DAQFVDVHSDTDALGYKEPLGNDFYPNGG	Į.						
LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHINAKVHYNYANFLKDQGRNKEATY HYRTALINNIKAWDYJCWRFRKTLTDLP  167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLLIGKELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR VEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFENECIVKINNVDLVDKTFA QAQDVFRGAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTICETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHSDTDALGYKEPLGNDFYPNGG	j						
GYQTLPHNAKVIYNYANFLKDQGRNKEAIY							
HYRTALNNKAWDYLCWRFRKTLTDLP	ì						
167 1517 A 2025 696 71 AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIMRFRKKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPKVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH TVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSQKA  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	1						
ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLLIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAFRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH TVATHLTKILNSDEHAVVISSAKTILCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDPYPNGG	167	1517	A	2025	696	71	AAASAASSI TVTI GPI ASACSUSII PPSGPGA
PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKPLLGKDLV RVEATTVIEKTESWPRIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVKAPRLLQVSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				] ====	0,0		
GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTILGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRILQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNIGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGFIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADA VASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVHISDTDALGYKEPLGNIDFYPNGG	Į.						
CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMFRKKNFKKKRIV TTPQTVLRINSIELAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVRAPRLLQVQSQVS RKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	1						
RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL  168 1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IFSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
1518 A 2046 2 366 HLQVAARVFMPLQAVVSAPKPLKGQAQAPQ RLQGARRVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR RVQAQTQTSEPQDLDQVPEEFQGQDQVLR 169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				] .			
1518 A 2046 2 366 HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR 169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH 170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	168	1518	A	2046	2	366	
PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	1		Ì	[			
KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR  169 1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	1			[			
1519 A 2049 1 945 QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				<u> </u>			
TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	169	1519	Ā	2049	1	945	
EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				<b> </b>			
QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				1 1			EDNSRSKREGLFHENECIVKINNVDLVDKTFA
VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				}			VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH  170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							
170 1520 A 2050 363 1 PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA 171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							ETASLVIARQEGHFLPRELVMFRSQSH
VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	170	1520	Α	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA  171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				( l			VAKVEKTYDKTLENAVVADAVASKCSVLNE
171 1521 A 2055 139 675 IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG							KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG				<u>[</u>	<u> </u>		ESSSEESLGESKEQLGDDVTKPSSQKA
DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG	171	1521	A	2055	139	675	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS
I DODOODETT CODOURS OF THE TOTAL TOT							DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
	L		L				LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

						(1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	ł	ł	peptide	l	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	ł		ł	}		RESCTITAYPCDSYQDYRNGKCVSCGTSQKE
	1					SCPLLGYYADNWKDHLRGKDPPMTKAFFDT
				L		AEESPFCMYHYFVDIITWNKNVR
172	1522	Α	2056	3	361	LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK
						MEWKVKIRSDGTRYITKRPVRDRILKERALKI
	,	1	}	}	ļ	KEERSGLTTDDDTMSEMKMGRYWSKEERKQ
						HLVRGKEQRRRREFMMRIRLKCLKES
173	1523	Α	2060	1	387	GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL
			l.			QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG
			ļ	1		AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA
		Ì				KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA
ļ	}		<u> </u>			G
174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE
		l				RRLKEEEEARLKYEKEEMERLEIQRIEKEKW
		1				HRLEAKDLERRNEELEELYLLERCFPEAEKLK
ĺ		ì	1	1	l	QETKLLSQWKHYIQCDGSPDPSVAQEMNT
175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV
1	ł	l	Ì		1	AESSTVGWLCALFRVTHVGVGATGHGVVCG
1		l				RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ
1	ł	1	1		1	RLQFPYLEPGHELPATTLLAFLAAV
176	1526	A	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP
	į	i				FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT
		İ	1		!	GIHSVYTVYQGHEIMFHVSTMLPYSKENKQQ
	1					VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS
		1		<u> </u>		HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG
	ŀ		1	1		PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET
j	ļ	ļ		}	1	PCI
177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC
	1				Ì	CDGAWLAWACWVFGNDFPSPASAACSALLG
	l	1	1	1		CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR
ļ	1	1				RAVSVPLTLAETVASLWPALQELARCGNLAC
İ	İ		1			RSDLQ
178	1528	A	2104	2	409	ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ
	Į.	1		1		PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF
	1		1			PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL
		1				LILQVREADAAVYYCAVEVPNTDKLIFGTGT
-		1	1	f		RLQVFPNIQNPD
179	1529	A	2111	1	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL
	1	1				MRTESHTGLKKGGNANLVFMLKRNTEPKKG
1	1	1	}	1		SYHFDLERLRAAHILFEREQEHLAPGGISMPL
		1	1			PPPLPLPACLG
180	1530	A	2116	3	366	TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE
1 - 3 - 3		1	1	1		LCRVMFDALEQKWKQTEQADLINELYQGKL
	1		1	1	1	KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS
1		ŀ				QAFASVVCTFHLTACVSLHRIHNSTVV
181	1531	A	2117	2	386	YGLGAHFGRLFIQAGINENDFYDGAWCAGR
			1	-	1	NDLQQWIEVDARRLTRFTGVITQGRNSLWLS
1		1	1			DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE
				1		GNSEKEIPVLNELPVPMVARYIRINPQSWFDN
1		{	1		(	GSICI
182	1532	A	2123	1	493	RTKTDVYILNLAVADLLLLFTLPFWAVNAVH
102	1332	A	2123	1.	1	GWVLGKIMCKITSALYTLNFVSGMQFLACISI
1	1	1		1		DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI
i	1	1	1	1		LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL
J	į.	1		1	1	IQMLEICIGFVVPFLIMGVCYFITARTLMKMP
		1	1	1	Ī	NIKIS
183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ
103	1223	1^	2170	1	1 301	RSGNYFVTMFDDYSATLPLLIVVILENIAVCF
]	1	1	Į	1	1	VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI
	<del></del>	ــــــــــــــــــــــــــــــــــــــ		<del></del>		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	"00	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	vence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uonoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence		ļ	/14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
		ľ		sequence		nucleotide insertion
		<del> </del>	<del> </del>	Sequence		SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS
						EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
					}	RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND
104	133.	**		١	330	NPPVFTRASYRVTVPEDTPVGAELLHVEASD
			1			ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR
		i				LAHALDCETQARHQLVVQAADPAGAHFALA
		1	[			PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG
				ļ		TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW
103	1555	^	1 2131	1 -	071	PKNNFNGSLVQASYQHEELRREVIMLACSFG
		1		1		NKHCHQQASTLISDWISSNRNRIPLNVRDIVY
			1	]		CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL
						LEALTCSDDRNLLNRLLNLSLNSEVVLDODAI
	ł			ł		DVIIHVARNPHGRDLAWKFFRDKWKILNTRI
	į			ļ		RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH
	ļ	ļ	1	į		E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYTVHPLWETWA
100	1550	**	1 2133	1	100	HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP
		1				LDEQNRDWQGLLENLHVELTLDEEDSEGPEK
	}	ł		ļ		EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC
				1	j .	AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR
		}		:	! · · <del>-</del>	PGAVAYTCNPSTLGGWGGWITRSGVRDQPG
	ĺ			ŀ		OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA
100				1	100	CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP
		<b>.</b>				AVVVPYMMVLQENGYGVEEGIPTLLMAASS
	<u> </u>			1	l	MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS
	1	ł		ł		LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG
						HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL
		l		1		QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF
				Ì		QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE
		I				LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV
		ł				AIVVSSLDW
190	1540	A	2179	64	399	MRLNQNTLLLESFGXXRPYTSEHAPTYHOW
	[	1	[		ĺ	MKADELLRWTTSEPLTLEHEYAMQRTWLED
	1		1	1		AYECTFIVLDAEKRHAQPGATEESCMVGDVN
				1		LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR
		1	1			LSYVLFIQERDVHKGMFATNVTENVLNSSRV
		ľ				QEAIAEVAAELNPDGSAQQQSKAVNKVKKK
	1	ſ	1	1		AKRILQEMVATVSPAMIRLTGWVLLKLFNSF
	l			1		FWNIQIHKGQLEMVKAATETNLPLLFLPVHR
		1			1	SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR
	1		1	I		DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS
						YTHSKGIMHRDVKPLNILCNSPRNKVILADW
	l	l	1			GLAEFYHPMRKYSVHVATRYYKSPEILLDYE
	1	1		1		YYDYSLDIWAVGVILLELLTLKLHVFEGGDN
1	1	1		1	[	EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS
]	1	)	]	I	]	NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE
	1			1		LPVPMGARYIRINPQSWFDNGSICMRMEILGC
		I	1	1	1	PLPDPNNY
<u> </u>	1	1		i i	ľ	LEIDINNI
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Lonce		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Gence	}		117	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ĺ	İ	•	peptide	sequence	/=possible nucleotide deletion, \=possible
ł -				sequence		nucleotide insertion
		<b> </b>	ļ	sequence		KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE
1	<b>!</b>	1	ì	1	[	MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW
	ì					MILKEDORQUAKIPDOTVENICIDQETTIMW
ļ		}	l .			MKHGPSPGVRAEKETILCYSDKTEMNRHHY
İ		1	Ì	Ì	ţ	ALYVHNCRLVFLLRKDFDQADTFRPAEFHW
ł						KLDQQALAKVDGQPGKSITRQLQEMPVTIQG
			1			ISLKPS
196	1546	Α	2256	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP
			ļ	l		GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD
ļ					1	NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF
}	•	1	1			KALEESGALLESGTYDGNFYGTPKPPAEPSPF
	1	ļ			,	OPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI
197	1347	Ι Δ	2239	133	3,74	WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV
	l		l .			VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI
1	]	}	1		1	KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP
	1	ļ	1			DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL
	[	!				
		<u> </u>	<u> </u>	<del></del>		HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF
1		ļ	1		1	FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT
		1	1			LMTRKICLQMMMASWMVGFLFSLCIIVTVFN
ì	ſ	İ	1	1		LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM
1	1	1				AIFVLSA
199	1549	A	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL
	1		1	l		RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI
	1		1	i		KRLAFCHSNVLSHSYCVHQDVMKLAYADTL
ł		}				PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF
200	1330	1.		-		SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD
			1			ILKQKAHQLASMQVQAYNGGNANPRPANNE
1	1	1	1		ļ	EEEDEEDEYDYDYESLSDDNILEDRPENKSCH
	}		1	1		DOLOFEYKEEM
-	l	<del>                                     </del>	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM
201	1551	Α	2330	13	312	TEELEALRSSSLGSRTLDPLWKVRRSQKLDM
	[		1			SARLELQSALEAEIRAKQLVQEELRKVKDAN
			1	Ţ		
1			ı	1		LTLESKLKDSEAKNRELLEEMEILKKKMEEK
ĺ			[	1	1	FRADTGKLMLCDSALFEYKYFSNECFYFLFD
	]	J		<u> </u>	<u> </u>	LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL
	[	1		1		NRVLERLAGGATRDSAASDILLDDIVLTHSLF
	]			1	}	LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA
			1	1		CLAMLLHFLDTYQGLLQEEEGAGHIKDLYL
			1	1	1	LIMKDESLYQGLREDTLRLHQLVETVELKIPE
	· ·		1	l		ENOPPSKOVKPLFRHFRRIDSCLQTRVAFRGS
		1	1	1	ĺ	DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV
1	1			1.	[	TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ
	1		1			PTEDCVFTALGINSHLFACTRDSYEALVPLPE
1	1	1	1	1		EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR
	i			1		CVHELEFVDYVFHGE
202	1553	A	2361	12	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH
203	ددد! ا	A	2301	1	1 703	GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD
	1	1		1	1	WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL
		1	1		1 .	
]	1	1		1	1	LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ
				1	<u> </u>	LGRITGLDP
204	1554	Α	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ
	ŀ	1		ļ.		AGSRLGAMRRCAREMDATPMPPAPSCPSERV
1					<u> </u>	T
205	1555	Α	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN
1	1	1				HGQDRRPHAPGRPARGKVQEGSARPPSAVAC
		1	1			EDCSCR
<u> </u>			J	<del></del>		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG MGANASTSPRCWDLSSGNKKWIIQVPILASIV ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	A	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT HTIMVLDKENTLS
208	1558	Α	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP NASNLDKVLTDIKADKDQANDGLSSALLILY LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG
211	1561	A	2431	1	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG
212	1562	A	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ
213	1563	A	2445	1	1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGERPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	peptide seq-		in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			į	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						AMVDĞRWLDÇVSHHDQLFRDEYALYRPLQV
215	1565	A	2464	3	2932	LFSVYCQLECSKLIL GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ
213	1303	A	2404	3	2932	HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
						CSLISGOHGPGESVSYEDDDIPAPASLLHVNA
						AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
						QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
						TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
					•	STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
						ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
				1		TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH
			)	ļ		ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC
						TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
		İ	1			ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
1						TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
						STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
ĺ						TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
[					ļ	STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
1		1	l	•	ĺ	POLITCAVSIHSTVPGVRISSRTPDLTCAVSTHS
ŀ						TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS
İ			İ			TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
		Ì		Ī		PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
						TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
i	1	•				PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
İ						TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
	Į					PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
ļ	]	}	ļ	Į		STVPGVHISSRTPDLTCAVSIHATVPSVHISSC
016			0.455	<u> </u>	414	TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	A	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
1		ĺ	l		ĺ	LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
	]					RVLARLDRDFLVHSSPHVALSHVDARSYHLL
]		]	J		ļ	VRDVSKENSGYYY
217	1567	A	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
		]	1			MQVLVCQHECVRELATRPGRLSPIENFLPLHY
1		1	l	1	1	DYLQFAYYRVGEYVKALECAKAYLLCHPDD
		-	1	}		EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
<u> </u>		<u> </u>				VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
	1			1	1	SANLLQLVRSSGDIQEGDLVEVVLSASATFED LOIRPHALTVHSYRAP
210	1569	-	2489	3	428	LQIRPHALI VHSYKAP   SSRLVLLAGAAALASGSQGDREPVYRDCVLQ
219	1309	A	2407	'	1420	CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
	ĺ	1				DCKYECMWVTVGLYLQEGHKVPQFHGKWP
	1			i	1	FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
1					1	FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
1		ì		1		APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
					-	HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
	]		]	1		MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
[	1.	[-			[	KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
		1	1	1	1	KEDLEKADALTFQTLNFLKSSLENTSIELHVL
	1	1	1	1	1	SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS
		i	1	]		LEGANAEIGLKENLQNTNALNSQTQAFIKSS
L		<del></del>	<del></del>	<del></del>		Provide Introduction of th

NO. of NO. of and ID NO. nucleotide editide sequence where the sequence of the	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq-			1				
Sequence						***************************************	
1949    1949	1		1	1			
1914   ng to first amino acid residue of peptide pep			j				
amino acid residue of peptide residue of peptide sequence   T-Threonine, V-Valine, W-Typotopalan, Y-Typotopal	, ·		Í			ı	
	donoc	] .	}	] ***			T=Threonine V=Valine W=Tryptophen
peptide		i					
		1	1			Sequence	
PDNTSAEIGPLERAGDEHIVLKRDLKM	1						
VTAGTOKANGRLDOTDTOJOVPKSEMENNN THAOJOVINGHMANASRIDITKOGMKNA SALTSOTOMLDSNLQKASAEJQRLGOLENT THAOT THAOJOVINGHMANASRIDITKOGMKNA SALTSOTOMLDSNLQKASAEJQRLGOLENT KALTMEIQGEGSRLKTHAVVTISQEQUERTO TEDTRHLSRKFKDWAYGPYYSSLYDLSSLD TCGEEASVLEILVYTNSKERNRHEMLAVEPINE LLKDKWRKFGAVSFYINVSYLCAMVIFILT AYYQPLEGTPYPYRTTYDVILKAGEVTLFFT GVLIPFTTY TOVILKAGEVTLFFT GVLIPFTTY TOVILKAGEVTLFT GVLIPFTTY AYYQPLEGTPYPYRTTYDVILKAGEVTLFT GVLIPFTTY AYYQPLEGTPYPYRTTYDVILKAGEVTLFT GVLIPFTTY AYYQPLEGTPYPYRTTYDVILKAGESTYSK QLENELGLKQKQISYSGYCSIBHQQEITKL KTDLEKKS	ļ	<del></del>	<del></del>	<del> </del>	sequence		
TLNAQIQVINGHMKNASREQITIKQGMKNAS   SALTSQTQMLDSNIQKASABIGNERGDLENT   KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ		1			l		
SALTSQTOMLOSNLQKASABIQRLRGDLENT   KALTMEQQEQ, RTQ   KAT		i					
		1	ĺ	İ			
1571	l	1			Ì		
TDEDTRILSKEKDWAYGPVYSSLYDLSSJ   TCGEASVLEILVYPNSKIENRHEMLAVPPINE     LLRDKWRKFGAVSFYINVVSYLCAMVIPILT     AYQELEGIPYPYPTRITVOYLLAGWIPILT     AYQELEGIPYPYPTRITVOYLAGWIPILT     AYQELEGIPYPYPTRITVOYLAGWIPILT     AYQELEGILKQKGDLRRTER     AKKELEVHTBALAABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLERTER     AKKELEVHTBALAABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLERTER     AKKELEVHTBALAABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLERTER     AKKELEVHTBALAABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AKKELEVHTBALABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AKKELEVHTBALABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AKKELEVHTBALABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AKKELEVHTBALABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AKKELEVHTBALABASKDRKLREQSERTYSK     CLENELGILKQKGISYSPOVCSIENGLEGIKTER     AVAPYKESI-PVSGEESQLTPEKSPKFPDSQNQ     LTHCSLSAA     CHALLES     CHALLE		1571	-	0501	<u> </u>	500	
TCGEEASVLEIL VYNSKIENRHEMLAVEPINE	221	13/1	A	2501	3	300	
LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT   AYPQHEGTPYPTTTVDYLRLAGEVITLFT   GVLFFFTN   A	1	1			[		
AYYQELEGTPPYPYRTTVDYLRIAGEVITLFT	Ì	Ì		[		1	l .
222   1572   A   2508   3   395	1						1
222		Į.					
ARHYRÖKEEEVÖLLYMÖK VESLAGEL RETTER		l			l		
AKKELEVHTEALAAEASKORKLEGSEHTYSE	222	1572	A.	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
AKKELEVHTEALAAEASKORKLEGSEHTYSE		ļ.	ŀ				ARHVRDKEEEVDLVMQKVESLRQELRRTER
		1	l		j	ļ	
						i	QLENELEGLKQKQISYSPGVCSIEHQQEITKL
MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS		1					
MYDERTDYLTKSI.KEKTPFSHCDQAVLQCS	223	1573	A	2544	2	412	NDPAIISNESAAVVHTIVNETLESMESLEVEK
EASSNKDMFADRLSKSIKHISIDKSKSVIPNID   KNAVYKESLPVSGESQLTPEKSPKFPDSQNQ   LTHCSLSAA   1574   A   2552   401   1   GASLCFISTAFTVLTICIDSCRFSYPERPIIFLSM   CYNIVSATVRLITVGRERISCDFEEAAEPVLI   QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL   TWFLAAGLKWGHEAIEMHSSYFHLAAWAIPA   VK		ļ					MVDERTDYLTKSLKEKTPPFSHCDOAVLOCS
SKNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ   LTHCSLSAA     224		ŀ	ł	ł	-	ļ	
LTHCSLSAA				ŀ			
224							
CYNIYSIA YIVRLTVGRERISCDFEEAAEPVLI QEGLKNTGCAIIFILLMYFFGMASSIWWVILTL TWFLAAGLK WGHEAIEMHSSYFHIAAW AIPA VK  225 1575 A 2563 724 1 MSARKERREKGEEGGEKDGDEDEKEEEKE GLGEEEKEAGKKKKKQEKEKEKGAVYSR VARICKNDMGGSQRVLEKHWTSFLKARINC SVPGDSFFYFDVLQSITDUQNGPTVVGVFTT QLNSIPGSAVCAFSMDDIEK VFKGRFKEQKTP DSVWTAVPEDK VPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLANDSAVPPIADEPWFT KTRVSYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFIFIKV AGAACDRGMSLACCAVTRKANRRTYTMG VALEFCSLPQTRRHNFEIGAEEMEIGMGHGE RGVIREKMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDPEKEEGILGSILFSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACFAQTVHYRPINLSSESHIVNVS LADLRGGNRPTNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLFPNSKTRESIGSVTP STHENKTLGFFAGEKERSMRDDTMWQLYEW QQRFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRHRHPKYK  228 1578 A 2583 3 330 LPFLGIGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRGGVVTVLVGNVQTAVE KIHAKVRGTWPFISPEVRREGGLPQTGRELLD PTMGIKPFLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCRPGYIGSNCOTALCDPDCKNH	224	1574	A	2552	401	1	
QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL TWFLAAGLK WGHEAIEMHSSYFHIAAWAIPA VK				2332	'''	•	
225 1575 A 2563 724 1 MSARKERREKGEEGGGEKDGDEDEKEEKE GLGEEEEKEAGKKKKKOEEKEKGAVYSR VARICKNDMGGSQRVLEKHWTSFLKARLNC SVPGDSFFYPVLQSITDIQNGFPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRIVLTTDDISSSPIEDRDGRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRINFFIGAEEMEIGMGIHGE RÖVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYSGSACPAVPHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRINMFS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGFGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRISTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHFKVK  228 1578 A 2583 3 330 LPFLGIGSVLPQGMVMASPEMIPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTVPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWWAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH						İ	
VK   MSARKERREKGEEEGGEKDGDEDEKEEEKE   GAVYSR   GLGEEEKEAGKKKKKQEEKEKEKGAVYSR   VARICKNDMGGSQRVLEKHWTSFLKARLNC   SVPGDSFFYFDVLQSTIDIQINGIPTVVGVFTT   QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP   DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK   TSIDFPDETLSFIKSIPLMDSAVPPIADEPWFT   KTRVRYRLTAISVDHSAGPYH							
1575   A   2563   724   1   MSARKERREKGEEGEGEKDGDEDEKEEKE GLGEEEKAGKKKQEKEKEKGAYYSR VARICKNDMGGSQRVLEKHWTSFLKARLNC SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSSWTAVPEKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPHADEPWFT KTRVRYRLTAISVDHSAGPYH   226   1576   A   2571   449   3   EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRIVLITIDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGHGE RGVIREKMPADAIVDHIMDRIFS   227   1577   A   2575   3   1197   VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKTSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDQDPFLTKINSVKLINSL PSEYESGSACPAQTVHYRPINLSSSENKUNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRCHEEETRGVISYQTLPRNMFS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDIMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTISPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK   228   1578   A   2583   3   330   LPFLGLGSVLPQGMVMASPEMNTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTVFFISPEVREGGLPQTGRELLD PTMGIKPHLWWVAA   DDKNAQGIKRHVKFTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	İ	İ		ĺ		ĺ	
GLGEEEKŁAGKKKKQEKŁEKGAVYSR VARICKNDMGSQRYLEKHWTSFLKARLNC SVPGDSFFYFDVLQSITDIIQNGIPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRIVLITIDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNIRVLIKPEIQNNGOKKEMSKIEE KKALEAEKVGFQKDGQDRPLTKINSVYLINSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVIP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGIGSULPGGMVMASFEMNTICSVFEA HIVLLFHATTERGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKFTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPPCKNH	225	1575	<del> </del>	2562	724	1	
VARICKNDMGGSQRVLEKHWTSFLKARLNC SVPGDSFFYFDVLQSITDIIQNGIPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRISTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWVVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR	223	13/3	^	2303	124	1	
SVPGDSFTYFDVLQSITDIIQINGIPTVVGVFTT QLNSIPGSAVCAFSMIDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGDVMNFEMAAEMAQEVAIP TRTVLTIDDISSSPIEDRDGRRGVAGNFPIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHEKTLGPGAEEKRRSMRDDTMVQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR	1	l	ĺ	1		}	
QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKKPRPGCCAKHGLAEAYK TSIDFPDETLSPIKSIPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRITVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRFESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPPGMWASPEMPTICSVFEA HIVLLFHATTIFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR	ľ	l					
DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSPIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETIRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVITP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWVVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCCTALCDPDCKNH	1		ĺ				
TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH  226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRIVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRINFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYSEGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFIGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTTRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWVVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCCTALCDPDCKNH		1	l			:	
KTRVRYRLTAISVDHSAGPYH	1	1					
226		1	i		{		
TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKAALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	206	1.00			4.0		
AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIPS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNINQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYJGSNCOTALCDPDCKNH	226	1576	Α	2571	449	3	
VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLIFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	ł	ł	l	l	ļ		
RGVIREKMMPADAIVDHIMDRIFS  227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	1	1					
227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	ļ	ļ		1			VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAA YQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH			L				RGVIREKMMPADAIVDHIMDRIFS
ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	227	1577	A	2575	3	1197	
TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAA YQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH				!	ł		EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH		1	}	1	ļ		ELWMKAMLDAALVQTEPVKRVDKITSENAP
KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	•			1			TKETNNIPNHRVLIKPEIONNOKNKEMSKIEE
PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAA YQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	]	l		ĺ	}		KKALEAEKYGFOKDGODRPLTKINSVKLNSL
LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	i		l				PSEYESGSACPAOTVHYRPINLSSSENKIVNVS
LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	1	-	l	]		i	
HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAA YQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	ļ	1		1	ļ		
STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAA YQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	1	1	]	1			
QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH			ļ	1			
MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK  228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	1			1	1		
228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	[	1	1	Í	1		
228   1578   A   2583   3   330   LPFLGLGSVLPQGMVMASPEMNPTICSVFEA   HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE   KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD   PTMGIKPHLWWVAA   229   1579   A   2589   1   448   DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR   ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	i		1	1			ORGDALIDEBHB THTDRAK
HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA  229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	228	1578	<b> </b>	2583	3	330	I PEL CI COM POCHMAN A CREATING COLUMN
KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD   PTMGIKPHLWWVAA		15,0	l **	1 2303	_		HIVE I FILATTED DO CONTROL VON THOM OF A TOP
229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	j		j	]	j		KINAKADCAMDERDEADARCO BOZOBELO
229 1579 A 2589 1 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	[.	[	[	[		·	DEMONDER WITHSTEVEREGGLPQTGRELLD
ECVAPNICKCKPGYIGSNCOTALCDPDCKNH	220	1570	<del> </del>	2500	<del></del>	440	
GKCIKPNICQCLPGHGGATCDEEHCNPPCQH	229	פוכו	A .	4369	*	448	
	ĺ	İ	İ	ĺ			ECVAPNICACKPGYIGSNCQTALCDPDCKNH
	L	L	L	L	L	L	GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

	1 000 00		1.550			Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]		}	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	Soquence	/=possible nucleotide deletion, \=possible
1	}		Ì	1	1	nucleotide insertion
	<del> </del>		<del> </del>	sequence		GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
		l	<b> </b>	1	İ	
	L					ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
	1	Ì	1			LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
1	}	l	ł	l		WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
232	1502	1 **	1 23,0	1 *	1	GRARRTPTCEPATPLCCRRDHYVNFQELGW
			l .		İ	RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
1	j	ł	]			
		ļ	Ì			FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
			1			LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
1		ļ				YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
	<u>[</u>		Į.			NSIPYWERIT
234	1584	A	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
254	1307	12	2017	1 .,0	555	DGSKSSDDQKIISYLWEKTQ
	1505	<u> </u>	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
235	1585	Α	2616	2	890	
1	ł	1				AKFLNVEAAMVFGMGFATNSMNIPALVGKG
j	1	į		1	ļ	CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
			ŀ	1		QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
	1	ĺ	1			YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
1	1	{				GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
	1	1	Į.	į		FGASGGYIAGRKARILSPPACLVPNTGSHSLH
Ì	ł	1	1	1	1	RLTRDLOMNEAMVALVTDRLOGWNSGEGN
1	İ		Į.	İ		WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
	j	ļ			l	AEQIIRSLKLIMGLDGTTQ
<del></del>	1.505	<del> </del>		<del> </del>	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
236	1586	A	2621	1	392	NISSTPACESSPARPSLERICSCRIPSNELLFLAS
		İ	1	ļ		ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
i	1	1	1	1		WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
	1	1	Į.	ļ		KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
	1	ļ	1	)	}	A
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
		1				WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
1	l		1	1		GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
		1				LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
	ł	ì	i	İ		
L	1.555	<del>↓</del>	1	<b> </b>	1104	QDPEPPNV
238	1588	A	2631	1	1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
1	1	1		1	1	ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
1	1	(				WQPCSRTCGGGVQKREVLCKQRMADGSFLE
	1	1	1	1		LPETFCSASKPACQQACKKDDCPSEWLLSDW
1	1	1		1	1	TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
1	1	1		1		TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
Į	1	1	1	}	1	AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
ì		'	1			VVLRCPARRVRKPLITWEKDGQHLISSTHVT
1	1	Į.				
ı	1	I	(	1		VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
1	1					VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
1	1	l	1	1		KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
1		1	1			RYDDLVSRLLEQGAPCSSSKKKN
239	1589	A	2636	1	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
1		1		1		TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
1	1		1	į.	-	SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
1			1	1		
}	1	1	1	1		TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
1		i			1	LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
J	]	J	}	}		GRLNCDPTFELEEMILESKPLHKKKKRLAKK
	+ .	1 -		-		EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE
1		l		1	l	KVNRDCI
240	1590	A	2639	389	3	ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
				1		
240		1	ļ.	i	i	LWONFAREIEEHVFTLYSKNIKKYKTCIRSKV
240						LWQNFAREIEEHVFTLYSKNIKKYKTCIRSKV ANLKNPRNSHLQQNLLSGTTSPREFAEMTVM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EMANKELKQLRASYTESCIQEHYLPQVIDGTL
241	1591	A	2640	392	3	Y IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	A	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1		MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	А	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	A	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
	1001	Α	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

(CEO II)	CEOTO	1364	LCEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	1	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	peptide		in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	seq-	ļ	09/496		to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	•			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ	1	ł		peptide	sequence	/=possible nucleotide deletion, \=possible
1						nucleotide insertion
<b> </b> -		<del> </del>	<del> </del>	sequence		OKRKKKAPDHSSGRKEELVTTHTVDKLETKK
]						PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
)		l	ł	ł	1	RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
	1			ļ		GAGDCL
252	1602	A	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG
232	1002	Α	2077	421	•	TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
}	]	)	ļ	1		VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
		1				MNVTGPQIRDGGVYRCTARNLVGSAEYQARI
ļ					}	NVRGPPSIRAMRNIT
253	1603	A	2698	65	401	ACCOWRTLIPAKSTTVSCTISTPHHPFRGSYS
233	1005	Ι Δ	2096	63	1 401	FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
1						PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
				,	ļ	AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
234	1004	^	2099	430	301	RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
233	1003	^	2700	1 1	042	AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
						DKETLENMMQRHEEEAHEKGKILSEQKAMIN
·		}	ļ	-		AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
1		l	ļ			NMKAQEEMISELRQQKFYLETQAGKLEAQN
		1	İ	1		RKLEEQLEKISHQDHSDKNRLLELETRLREVS
[		[	1		(	LEHEEQKLELKRQLTELQLSLQERESQLTALQ
				1		AARAALESQLRQAKTELEETTAEAEEEIQALT
		[	1	[	1	VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
250	1000	1.	2,01	1	1 103	LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL
1	1	<u> </u>			į	YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA
	İ	(	ľ	l		FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
	1		l			PYEAARMFFEGLR
257	1607	A	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
	1 2007			-	""	FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
}	ł	i	ł	1	l	NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
	1		1			DLPVOARRLVDLMKNDTRIHFQEDWKIITLFI
	ļ	1	ł	i	1	GGNDL
258	1608	A	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
	1	}		ļ <sup>-</sup>	}	MTRKELLTVYSSEDGSEEFETIVLKALVKACG
Ì				1		SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
			İ			QWRSFHLEASLMDALLNDRPEFVRLLISHGLS
	ļ	1	<b>\</b>	ļ		LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
{	Ì	ľ	1	ì	1	SAGTKAPALKGGAAELRPPDVGHVLRMLLG
İ	1	ĺ			1	KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
1	1	{	1	l	1	KATSPLSLDAGLGQAPWSDLLLWALLLNRA
1			1	1		QMAMYFWEMGSNAVSSALGACLLLRVMAR
	ł	l	1 .	l	l	LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
1		1		1		SSEVRAARLLLRRCPLWGDATCLQLAMQAD
1	}	}	]	1	1	ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
1	ļ	1	1	1	1	GGFSQREMVTGERSPSPEEEEEEEEEGFGERA
l	1	Ī	1			SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
]	]	}	1	1	1	PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
1	1	1			1	GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
[	1	1	1	1	1	LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN
1		ļ	J	J		IATVVKGAERASSMAGTKPYMAPEVFQVYM
I	1	1	ļ	I	1	DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
ļ	ŀ		1	1	1	HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
L	L		<u>L_</u>	<u></u>		LRK
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR
1	1	1	I	1	i	EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
•		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		J	]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ĺ	peptide		/=possible nucleotide deletion, \=possible
		<u> </u>	<u> </u>	sequence		nucleotide insertion
1					ĺ	RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
		ı				RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
i i		l	ľ			DIFVDRRVSALAVVNECGTHPQDERLGLGW
262	1612	Ā	2733	3	431	GLGEPGSEERLFPAAITSR
202	1012	1^	2/33	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
		İ				GRLVKLSLANNNLVGVHEDAFETLESLQVLE
		ļ			ļ	LNDNNLRSLSVAALAALPALRSLRLDGNPWL
!			}			CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM ESRRISLRACRRPASRV
263	1613	A	2736	2	343	
203	1015	) ^`	2/30	-	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
1		1				ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
			,	1		RLISPLYNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	RAMLKCLREGOPPPSYNWTRLDGPLPSGVRV
~~		''	2,30	~	~~~	DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
1 1		ł				DTVDVLDPPEDSGKQVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
1 -00	.0.5		2,32	-	300	LAAVETTVLVLIFAVSLLGNVCALVLVARRR
						RRGATACLVLNLFCADLLFISAIPLVLAVRWT
J I		}				EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
		1				SLER
266	1616	A	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
[	·				-	LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
1						V
267	1617	A	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
i i						HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
]						LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
						AVLLLLLLSLALGLVLAALGLFVHHRDSPL
1 1						VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
1						ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
						LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
						LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
						LSKNLSFSEFCFDVSY
270	1620	Α	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
ł (	i					VEQIAKAEETHSSLSQELQARLQTVTREKEEL
1						LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
	1			ì	ļ	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
						KAYDELRLQSEAFKKHSLDLLSKERELNGKL
271	160:		5505			RHLSP
271	1621	A	2795	1 -	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
				l		FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
] [				•		RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
1				ĺ		FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
1				1		VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
272	1622		2707	-	-522	SQRDDGKGALNIVEFLRKREELHQILSTVKQP
2/2	1622	Α	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
			}		]	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
						GYRKVVSNNCTDGVREQYTAKPQKCPGKAP
				ļ	}	RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
			j			VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
273	1623	<del>_</del>	2801	72	205	YQNXGIXRXTVQVDNSLGS
2/3	1023	Α	2001	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
		1		ļ		DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
[			. 1	Ì	_	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
274	1624	A	2805	168	320	AVSRKTELEEYH ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE
	-02-	*		-30	220	IFIARNGVVGETLTHCKRV
	1	1	1			II INTUIO A AODITIUCKKA

	T-22-2	<u> </u>	Laco			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	ł	in USSN	location	1	F=Pnenylalanine, O=Glycine, H=Histidine,   I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq- uence	•	09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	dence	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice	}	]	1 717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ	1	<b>\</b>		peptide		/=possible nucleotide deletion, \=possible
]	}	ļ		sequence		nucleotide insertion
275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
		ŀ				MGKIIFQ
276	1626	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
Ì		ļ	1			KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
			<u>i</u>			QVGPVRRNGEAGPG
277	1627	A	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
		1				LFISYLHTPKHKQHEVLQAMGSILGITGEEME
<u>'</u>		İ	1	İ		PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
				1	<b>\</b>	GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
050	1600			220	167	LPPHNSPGKIK
278	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
		ŀ				PVDONPRLV
279	1629	A	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
213	1029	1 ^	2022	1 342	1 '	TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLOS
						CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
					1	VPSORHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
		1	1	İ	ł	CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG
		1				VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
	ł				ł	QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
				•		NTTNMDEVPRPQALSGSSVVWVSGCVASRS
						VILSLTSG
282	1632	Α	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
	Ì			ļ		TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
		1		}	[	YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS YLKTLPPYYL
283	1633	A	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
203	1033	] ^	2055	102	140	MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
		1 .				PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
		i		,		SSEESAP
284	1634	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
		İ				DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
		l			ļ	KSLAETVLNFPLDKSLLLRCSNWDAETLTED
		1		l .		QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
			ļ			KR
285	1635	A	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
ļ			1	}		VCDRVSEDGINRQQAQEWCIKHGFELVELSP
206	1626	<del>  </del>	2046	107	278	EELPEEDGKCLCVRRKYGTYI
286	1636	A	2845	197	278 427	TAEDVLTVAYEHGVNLFDTAEVYAAGK FVAEVRREWAKYMEVHEKASFTNSELHRAM
287	1637	A	2631	*	421	NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
]	}	1		}		AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
				]		QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
	1			1	[	KVYLEQNLAAQDRVLCALT
288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
1		1				LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
			1.	1		TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
		1		}		KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
	<b>!</b>			<u> </u>	<u> </u>	EPGQELAETALHLAVRTADQTSLHLVE
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
		l		l	1	DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
						DSGLWRMHLMEGELPASMSGSCGACINGKL
ľ	1		1	1		YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
200	<del> </del>	<b> </b>	20/2	<del>                                     </del>	L	ITDFEGQPPTPRDKLSCWVYKDRLIYFG
290	1640	A	2868	] 1	378	FROGQLYKVFLHGSQGQVYHSQQVGPPGSAI
1	1			[		SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF PDCASCLQAQDPLCGWCVLQGRCTRKGQCG
L	<u> </u>	Ц		L	<u>L</u>	L D C V S C L C A C L L C A C

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	ļ -	/=possible nucleotide deletion, \=possible
				sequence	l	nucleotide insertion
						RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR
					<u></u>	QE
291	1641	Α	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
				i		PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
						GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
						AKVINAENAAHKSEKFRAMATRTRQEYLKD
						LA
292	1642	Α	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
000	1242	<u> </u>	2070	<u> </u>		PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	1	427	REKEEEVEEEEDKVVKETEKEAEQEKEEDSL
		ļ		j	ļ	GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
						LSPEKLTAENRYYCESCASLQDAEKVVELSQ
						GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
294	1644	A	2879	109	245	LRLPLAGGRGQAYDL
294	1044	A	28/9	109	243	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI IIVFVTGGVLG
295	1645	A	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
293	1043	^	2000	3	320	NNCVGEONHRFFCALHCKSKHFCIEFTLNTNF
						FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
	•				Ì	LSESISO
296	1646	Ā	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
290	1040	^	2072	209	303	RLQEFSQKMDQVRGHWPVST
297	1647	A	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
	10.,	**	20,5	1	127	KLYSTMGRFLRDRKNPACREMAVVLLANLA
						QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT
						QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
		ŀ		ł	ł	LLALAKVDDNHSEF
298	1648	A	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
					1	SGLLNASAQVNL
299	1649	A	2898	1	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL
						GYFQAYNVLILTMQASLPKVLRFCACAGMIY
		ļ	İ		İ	LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
		1			ŀ	VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI
						SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
						LQEF
300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
		l	1		}	TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
		1				SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ
		1			1	LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
						FSSYEKLDITVLDVNDNAPIF
301	1651	Α	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
		]	1		I	EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
		<u> </u>				KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	Α	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV
		1			ĺ	EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
		}	1	}	i	CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
		1	1			SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
202	1650	<u> </u>			<del></del>	LVNRQDRAHFM
303	1653	Α	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
204	1654	<u> </u>	2025	170		VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
205	1/22	<u> </u>	2022	125		NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	A	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
		1	1	]	1	DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
		1	1	i	1	PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE
306	1656	<del> </del>	2044	<del> </del>	200	LHAREWP
300	1020	Α	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
		1	İ	1		EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496 914	location correspondi	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
307	1657	A	2950	2	411	SKTKLSFKVSSSA NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL
307	1657	A	2930	2	411	PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT CTAENAVGRARRRVHLTILVLPVFTTLPGDRS
308	1658	A	2951	1	407	LRLGDRLWLR PTRPPRVRFDNEFDAESQRKRTTSVSKMERM
308	1038	A	2931	1	407	DSSLPEEEEDEDKEAINGSGNAENRERHSESS DWMKTVPSYNQTNSSMDFRNYMMRDETLEP LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP
			[			RLCKKAKAPEDC
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	Α	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ
						HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
						TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
311	1661	A	2963	3	465	LFLAVYAVATIMFCF MKPOMPGLGAPNGYGPGRAGVPGGPERR
311	1001	^	2903	٦	403	PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK
	E				•	PQKPGLRGTLKPQKSGHGHENGPWPGPCNA
	ł					RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ
1	i		1			KI .
312	1662	A	2967	3	405 .	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM
		}				HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
						ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI
			2000		420	DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW
			ļ			KVPHVQVKDVPNFEQLSPELEAALKKACTRD
						PSRWARFWHSGPGQVLTYLLLPCTLPFEYTYF RSRRLVVWLPDVPADLWWMQ
314	1664	A	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
31.	100.			'		LDALGRGVFVNASGLRLLDLSSNTLRALGRH
	ļ		ļ	ļ	]	DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
	}					SSNRM
315	1665	Ā	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
	1	•				QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV
						FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
						TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP
316	1666	A	2978	2	400	RSTTGTVHVAVLDLNDNT ELVVELVSAGKSGPERNTYEVQVVTGNVPKA
310		<b>  '</b> `	2776	[ -	100	GTDANVYLTIYGEEYGDTGERPLKKSDKSNK
						FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR
		ł	ł			AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE
317	1667	A	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR
	}	1			1	LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS
						DPSQYDASINLMNLQVSDTATYECRVKKTTM
<u> </u>					<u> </u>	ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ
				}		RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

456 TO	OFO ID	1 14-4	LOPO	Darding	I B. 19. 1	I
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of		noa	in in	nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide	l	USSN		location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	amino acid		Q=Glutamine, R=Arginine, S=Serine,
		i		residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
:				peptide	1	/=possible nucleotide deletion, \=possible
		ļ	<u> </u>	sequence	ļ	nucleotide insertion
1		ł	İ	1	1	STLALSHSAQVLASASGRSSTTAHCQIRVWD
		ŀ				VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
200	1650	<u> </u>	0000	(02		GDHDGRTLALWGTGHL
320	1670	A	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA
		)	1	]	1	PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
			1			ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
201	1.661		2001			TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	Α	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
		ŀ				GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
		i	i	1		GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
		<u> </u>				WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	Α	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
					•	LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
						DGAASPRNVGHNIYILAHQLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
				1	}	QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND
				1		ERVFGKRGF
324	1674	Α	3020	523	797 .	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
	,	ĺ	1	ĺ		YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG
		ļ				FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
						GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
			1	İ		FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
		ŀ				GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ
				1	1	EGDLVEVVLSASATFEDFQIRPHALTVHSYRA
		1	[		į	PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
					·	RC
328	1678	A	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
					***	LTGALNAHNKAAVDWGWQGLIAYGCHSLV
		l	ł	ł	ł	VVIDSITAQTLQVLEKHKADVVKVKWAREN
			1			YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
•						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
		ļ				HPPNYIVLWNADTGTKLWKKSYADNILSFSF
		ŀ				D
329	1679	Α	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
				**	,	GNIEYKKLVNPSQYRFEHLVTQMKWRLOEG
						RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
						HRMAEKVGADITVLREREVDYDSDMPRKITE
1		l		}	-	VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
						LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
İ		ļ				
		İ				TSSISFEILGFNSKGEVHGINGTQWGQTLRMG W
330	1680	A	3040	3	397	
220	1000	Λ.	3040	,	371	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
1	,					LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE
		<b>.</b>				WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ
		Ì				VVLTMTNMGPVDTATYYCAQFARGARGSN
-337	1/0:	<del> </del>	20.12		1400	WFDPWGQ
331	1681	Α	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK
Ì		•				MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK
ļ		1				GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
l	•		j			WGNHPKIETAAMDGTLRETLVQDNIQWPTG
ſ						LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
		1				VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
		[				FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
1					,	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
	!					QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN GGSCFLNARRQPKCRCQPRYTGDKCELDQC

			-	1 D - 12 - 1	D 2 2 1 . 1	LA de la companya (A – Alemina C–Custaina
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	İ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	Soquence	/=possible nucleotide deletion, \=possible
					İ	nucleotide insertion
[				sequence		
			i			WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
1		1	ļ			TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
l i	İ	ļ	]	ļ	}	FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
		1			l	CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS
i		1				GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT
•		}				MNSKMMPECQCPPHMTGPRCEEHVFSQQQP
1 1		1	}	1	l	GHIASILIP
	1.600	-	2045	1-3	052	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG
332	1682	A	3045	3	952	ITTISHTHIQVIKTICCOTTRACTIVIKQISEVO
						AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT
1	ŀ	İ	1			LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD
			)	}	}	MPKSPFKRRRSMNEIKNLQYLPRTSEPREVLF
[	[	ĺ	1		1	EDRTRAHADHVGQGFDWQSTAAVGVLKAV
1 1		!	1		ł	QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ
		1				LDLHEPPVSQCVQWVDEAKLNQMRREGIRY
	ļ	]	ļ	1	Į.	ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW
	ì	}	ł			HIRLKOYHPVVEATONTESNSNMDCGLTGKR
			1		ì	ELEVDSOCVRIKTESEEACTEIQLLTTASSSFP
	]	į.	ł	Ì		1
		<u> </u>				PASE
333	1683	Α	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL
(	l	1	Í	Í		YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT
	į .	1	l		ļ	ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG
	1	Į.	l .	İ	1	CSSKTWKVAPFVRAWWRP
334	1684	A	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD
334	1004	ļ ^	3033	"		EASGANDEIVQLRSEVDHLRREITEREMQLTS
	ĺ		Ì		1	OKOVRRVNKVVRSLEDF
		<del> </del>	-	·	1046	WDAWGDWSDCSRTCGGGASYSLRRCLTGR
335	1685	A	3054	2	846	
		1	1			NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA
	ĺ	İ	ı	ĺ	(	YNDVQYQGHYYEWLPRYNDPAAPCALKCH
1	1		ì	1		AQGQNLVVELAPKVLDGTRCNTDSLDMCISG
i	I	1		i		ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC
	]	1	j	}		RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
1				1		TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV
1	ļ			1		ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY
1	1	ł				TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT
	ļ	}			}	CGGG
	1.00	<del> </del>	120.55	+	242	
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA
}		1		1		LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG
1	ł	i	1	1	1	AWDWRLGSPACPHWGLHKLPRLWDPLSLYP
		1		1	1	VLCWGT
337	1687	A	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE
	1	1		1		QQQRRHPQHLHQQHHGDAAQHTRTWKLQT
1				1	1	DSNSWDEHVFELVLPKACMVGHVDFKFVLN
	1	}		1		SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF
1	1	1		i		LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
	1	1				
		1			1	SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI
i		1		1	1	CNGGMRPVVRLPSLKHQSNKGYSLASLLAK
1	1	1	L	1	1	VAAGKEKSSNVKNENTSGTRK
338	1688	A	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK
					1	DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL
	}	1	1			SDVNPRNTFGOLFCGSLDLFGILCVGLYRIIDE
1	i	)	1		1	EELNP
220	1/00	1	2062	226	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV
339	1689	Α	3063	236	302	
	L		<u> </u>	<u> </u>	<u></u>	PSSVTTMLSWV
340	1690	Α	3065	3	1249	DLWQFTPLHEAASKNRVEVCSLLLSYGADPT
	ŀ	-	1			LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL
1	i	1.	1			QAAREADVTRIKKHLSLEMVNFKHPQTHETA
1						
1			i .	1	ı	LHCAAASPYPKRKOICELLLRKGANINEKTKE
						LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL
						LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  IISLQGFTALQMGNENVQQLLQEGISLGNSEA DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE GRQSTPLHFAAGYNRVSVVEYLLQHGADVH AKDKGGLVPLHNACSYGHYEVAELLVKHGA VVNVADLWKFTPLHEAAAKGKYEICKLLLQ HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR GDAALLDAAKKGCLARVKKLSSPDNVNCRD TQGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI RTVKFLRSATIPVVELMDVQGERLDMEVGFD NRQAAFDMVCTMLEKRVRHKILYLGSKDDT RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI HLRMQLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE IDDNRVTAEEVDILLREGEKLAPVMAKTRILR AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE RDGLDGFITITGGKLMTYRLMAEWATDAVC RKLGNTRPCTTADLALPGSQEPAKVP
343	1693	A	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS LGASRAQVLWFVILPGALPEILTGLRIGLGVG WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV AHSKPSTRNILLLL
345	1695	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC VINQAAATGGDARQLVGYLVSQSGLPLDTSA LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL DRKALPLPELKAQAPGRAPKAGSETIIAAAFS SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	A	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH
349	1699	A	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR SALTYEPIFLILLEALRGVIHATPCRVSLSLWD GLDLP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		Į	914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ		peptide	sequence	/=possible nucleotide deletion, \=possible
			ļ			nucleotide insertion
<u> </u>	1500		2110	sequence		AQLAEVCPPQTLLTTNTSSISITALAAEIKNPER
352	1702	A	3110	341	2	VAGLHFFNPAPVMKLVEVVSGLATAAEIXIVE
1		ŀ		Ì		
1		ł				QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
						YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
						GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	Α	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS
L			L			FPFSNMTEVRGLVFLS
355	1705	Α	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
1	İ	ļ	i			EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
}	}		1	}	ì	ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
I		ļ			1	RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
	ŀ	1				LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
j	1	1	j	}	}	SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
	İ					DEPTN\HLGHGRAIEALGPCLQTISGVGVILVS
1					<b>,</b>	HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG
	1			1		AASREHARWQGTGLAPGTRVAVAPTCVQGL
1	]		ļ	}		POERSVCRPFFSSRWREGPVWALGAGAHGKP
	1		İ			RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
1 337	*/0/	Α	3124	1245		PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
	ĺ	ĺ	[	[	1	EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
	ł	1				LDRCVAFIGOARAEGRAVLVHCHAGVSRSV
	ŀ					AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
1	Ì	i	1	Ì	i	EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
			1			KVTEKYPELQNLPQELFAVDPTTVSQGLKDE
i	l	]				VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
1	l	ł	ł		ł	KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
	١.	· .	1			LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
Í	ĺ	1	[	Į	[	GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
260	1708	<del>                                     </del>	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
358	1708 .	A	3121	910	139	TPPP*LSGPGPOGAPVLGKLLPDPEETPAGKTP
		ľ	ĺ	1	l	LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
	]		1			
		i .				GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
	l	l		ł		PTANREINPGPAAAADTRSCWGHKRSWRGW RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
		i	ļ			KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
i		ļ	1			1
0.50	1500	<b> </b>	0:00	1	<u> </u>	VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
	1	ł	1	l	1	HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA
	<u> </u>	<u> </u>	<u> </u>			*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
		1	ļ	1	ł	AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
	· .	1	Į.	1	1	*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
		<u>l_</u> _	<u> </u>		<b></b>	QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
		1	1	1		AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
		1	1			VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
}	]	J		1		GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
		I				GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
			1	1		PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
1	1	1	1	Į.	İ	APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
		ĺ	1	[	1	GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
1	}	1	1	J	J	PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
		}				SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
		l	1.		<b>[</b>	GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
		1			I	F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR
1	1	1	1		1	OG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP
L	<u> </u>	ــــــــــــــــــــــــــــــــــــــ	ــــــــــــــــــــــــــــــــــــــ	1	ــــــــــــــــــــــــــــــــــــــ	1 40 . Dela Di diniquini Quini individir

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RNQSPLGNDTLSSGLPMGPRRQVWPLARVG GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
362	1712	À	3136	1270	274	DNL  RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQMV*GGSRDWEEGVEEQQVGNKFSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI SGASLCPVLSPPR*PLQPSSL
365	1715	A	3145	122	413	LLPYPSLFVFLRQCHFVTVRLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS/DK GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRK WLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718		3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVPMSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*1*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	Α .	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410	14101	RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSENPPI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	dence		914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
l danoe			717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	-	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						FEIR\MYDAQHQQVGSNKCRVNNAGCSSLCL
						ATPGSRQCACAEDQVLDADGVTCLANPSYVP
						PPQCQPGEFACANSRCIQERWKCDGDNDCLD
						NSDEAPALCHQHTCPSDRFKCENNRCIPNRW LCDGDNDCGNSEDESNATCSARTCPPNQFSC
						ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
						CFPLTQFTCNNGRCININWRCDNDNDCGDNS
						DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
						NDCGDYSDETHANCTNQATRPPGGCHTDEF
						QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
						GVTHVCDPSVKFGCKDSARCISKAWVCDGD
						NDCEDNSDEENCESLACRPPSHPCANNTSVC LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
						GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
						CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
						WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
						KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
		•				VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
					-	LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TLLAGDIEHPRAIALDPRDGILFWTDWDASLP
						RIEAASMSGAGRRTVHRETGSGGWPNGLTV
						DYLEKRILWIDARSDAIYSARYDGSGHMEVL
						RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
						KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
						QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
						ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
						QRVYWSDVRTQAIKRAFINGTGVETVVSADL
					· [	PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
						RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
				-		WTDGDNISMANMDGSNRTLLFSGQKGPVGL
				!		AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
						AMRSQLGKATALAIMGDKLWWADQVSEKM
		·		ĺ		GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
						MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
			.	ļ		GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
'						TIYWVDMGLSTISRAKRDQTWREDVVTNGIG
				l		RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
	' I			i	ľ	SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
						GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV
		l			ļ	LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
				-		RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
			į		ļ	QKGTNVCAVANGGCQQLCLYRGRGQRACA
]	ŀ	ł		Ì		CAHGMLAEDGASCREYAGYLLYSERTILKSI
			1			HLSDERNLNAPVQPFEDPEHMKNVIALAFDY
	J	j	j			RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
1 1		ļ				RHTVDQTRPGAFERETVITMSGDDHPRAFVL
			Ţ		}	DECQNLMFWTNWNEQHPSIMRAALSGANVL
		1	l	}		TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
]		]	į			RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
			ļ			WTDWVRRAVQRANKHVGSNMKLLRVDIPQ
		1	1		1	QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
	]	.			I	LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE
		1	1		į	KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
		ļ		ļ	ļ	GADDCGDG\$DEIPCNKTACGVGEFRCRDGTC
						IGNSSRCNQFVDCEDASDEMNCSATDCSSYF
					_ <del></del>	

Noc of Not of Not of Indian in a clotide cotted of USSN or unnec un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un entre un un un un un un un un un un un un un	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence cottde seq- uence corresponding to corresponding to corresponding to corresponding to periodic sequence s							D=Aspartic Acid, E=Glutamic Acid,
eoride uence    Sequence   1944   194		1					F=Phenylalanine, G=Glycine, H=Histidine,
peptide residue of peptide sequence pept	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence per seq	seq-	uence		09/496		to last amino	
residue of peptide sequence  Poptide sequence  RIGVS(VI) FOPCERTSLCYAPSW/CDGAND CDSVSERDE/DOVKRPRC/PLNYFACPSGRCIP MSWTCDKED/DCEHGEDE/HCNKPCSAQPE CONHRCISK,0WL/DGSDDC/DGSDEAAHCE GKTCGPSSSCCOTHWCVPERWLCDG/DBC/DA DGADESIAAGCL/NSTCD/DREFMCQNRQCIP HIFVC/DBIRDCADS DSESPECE/PT/GGSEF RCANGRCLSSRQ/WECDGSDE/CDGSDEAAHCE GKTCGPSSIC/CTGW/PRWCLOB/DBC/DA DGADESIAAGCL/NSTCD/DREFMCQNRQCIP HIFVC/DBIRDCADS DSESPECE/PT/GGSEF RCANGRCLSSRQ/WECDGENDC/DSDEAPHC NOW CONTROL OF THE CONTROL OF	uence			914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
peptide sequence    Possible   Includite insertion		1					
miclotide insertion  RIGVSQUI-FIPCERTSLCYAPSWYCDGAND  CGDYSDERDCPOVKRPRC/LNYFACPSGRCIP  MSWTCDKEDDCEHEGDETHCHKSCEAQPE  CQNHRCISKQWI-CDGSDDCGDGSDEAAHCE  GRTCGPSSSCCPTHCVPERWLCDGDKDCA  DGADESILAGCLYNSTCDDREFMCQNRQCIP  KIFIYCDHDRDCADGSDESPECCHYTCGPSEE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDQSDEAPHCE  RCANGRCLSSRQWECDGENDCIDGSDEAPHCE  RCANGRCLSSRQWECDGENDCIDGEDGENDCPE  RCANGRCLSSRQWECDGENDCIDGEDGENDCPE  RCANGRCSNLCLLSPGGGHCACAPTNYLGSDIP  IFALTLEEDYVYWTDWETKSINRAHKTTGTIN  RCALLSRPGALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPGALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPGALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPGALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWYTDWETKSINRAHKTTGTIN  RCALLSRPALVYWTDWETKSINRAHKTTGTIN  RCALLSRPALVYWTDWETKSINRAHKTTGTIN  RCALLSRPALVYWTDWETKSINRAHKTTGTIN  RCALLSRPALVYWTDWETKSINRAHKTTGTIN  RCALLSRPALVYWTDWECKSCONGCCTGTINCTIN  RCALLSRPALVYWTDWETKSINRAHKTTGTIN  RCALLSRPALVYMTDWETKSINRAHKTTGTIN  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETCHCAPPT  RCALLSRPALVYMTDWETC			[			sequence	
RIGVKGVLFQCERTSLCYAPSWCDGAND CDSYSBEDCPGVKRPCCLYNTACSGCUP MSWTCDKEDDCEHGEDETHCNKFCSEAQFE CONHACISKQWLCDGSDDCGDGSBCAAHCE GKTCGPSSSSCPGTHVCVPERWLCDGDKDCA DGADESIAGCLVNSTCDDREFMGVRQCIP KHFVCDHDRDCADGSDSPECYPTCGPSEF RCANGRCISSRQWLCDGHENCHORNQCIP KHFVCDHDRDCADGSDSPECYPTCGPSEF RCANGRCISSRQWLCDGHENCENDCDDSDEAPK NPHCTSPEHKCNASSGFLCSSGRCVARALLCN GDDDCGDSSDERGCHNECLSRLSGCSQDC EDLKIGFKCRCRPGFRLKDDGTTCADVDECS TTFCSGCNCTHTGSSYKLCVEGVAPRGGDP HSCKAVTDEEPPLIADREYLKLDGSNY TLLKQGLNNAVALDPEVREGMYYDTVTTC GSMRRMHINGSNVQVLHRTGLSNPDGLAV DWGGGNLWCDKGRGDTFVSKGMAYATVTUTTTC GSMRRMHINGSNVQVLHRTGLSNPDGLAV VSSGLREPALVDVQNQVLYNTDWGDHSL VSSGLREPALVDVQNQVLYNTDWGDHSL VSSGLREPALVDVQNQVLYNTDWGDHSL VSSGLREPALVDVQNQVLYNTDWGDHSL VSSGLREPALVDVQNQVLYNTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGDHSL VSSGLREPALVDVQNQVGTSVTDWGTSVTDAWG KTLLSTHRPANDLHVFHALRQPDVPNNPCK VSSGLREPALVDVGTSVTDWGTSNBALKTTGTN KTLLSTLHRPANDLHVFHALRQPDVPNNPCK VNNGCSSLLCLLSRGCHKCACPTPVLCSD GRTCVSCTLSQFVCXADKCGPWWCCDHFLCXDV VSSGLREPALVVDVQNGCSHALVDVGTSCTLCTN KTLLSTLHRPANDLHVFHALRQPDVPNNPCK VNNGCSSGLACKCGCQPCCCTCCTC VSNGCCNACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		}	} ·		• •		
CGDYSBERDÉPOVKRPRCZLATYFACPSGRCIP MSWTCDKEDDCEHGEDEHTCHKRSCEAQPE CQNHRCISKQWL_CDGSDDCGDGSDEAAHCE GRTCGPSSSSCPTHVVPERWLCDGDKDCA DGADESIAAGCL_YNSTCDDREFMCQNRQCIP KIFVCUBIRDCAADSBESPECEYPTGPSEE RCANGRCLSSRQWECDGENDCIDQSDEAPK PNHCTSPEHKCNASSOFLCSSGRCVAEALLCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRPGFHLEDDCRTCAADVDCS TTFCSQRCDTTHGSYKCLCVEGYAPRGDDP HSCKAVTDEEPFLIPANRYVLRKLALDGSNY TLLKQCLNNAVALDFDYREQMTYWTDVTTQ GSMIRMHENGSNYQVLHRTCLSNYDGLAY DWYGGNL-WCDKGRDTEVSKLKDAYRTVLV VSSGLREPFALVVDVQNGTLYWTDWGDBISL IGRIGMGDSSSSVTUPWTSTWTDWTTQLTYTT RYWADAREDYIEFASLDGSNRTVVLSQDIPH FALTLEPDVYYWTDWTSTSNRAKETTGIN KTLLISTLHEPMDLHYTHALRQDVPNIPCK VNNGGCSNLCLLSSGGGHKCACTINTYLGSD GRTCVSNCTASGFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCEFFKCRPQQQCSTGICTN PAFICDGDNDCQDNSDEANCDINVCLSQFK CTNTNRCPGFRCRQGQQCSGTGICTN PAFICDGDNDCQDNSDEANCDINVCLSQFK CTNTNRCPGFRCRGPQQQCSGTGICTN PAFICDGDNDCQDNSDEANCDINVCLSQFK KNRCYGRWQCDYDDDCONDSDEENCEDERDCE VTCAPNOFQCSITKRCPRVWCDRDNDCVD GSDEPANCTQMTCGPBFRCAGGQQCSGCGCACTINGCTPAW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNRCYGRWQCDYDDDCGNDSDESSCTPR RCSESEFSCANGRCIAGRWKCDGBHDCADGS DEKDCTFRCDMDGQCKSGGGCTPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNRCYGRWQCDYDDDCGNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGBHDCADGS DEKDCTFRCDMDGQCKSGGGCTPARWCDA DADCMGSGBEAAGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDGCNSGBCEPCCARFC CRYNRCYGRWQCDYDDGCONSDESSCTPR RCSESEFSCANGRCIAGRWKCDGBHDCADGS DEKDCTFRCDMDGQCKSGGGCTPATRYCDA ANSKCGGDCONSDENPECARFC CPPNRPPRCKNBRVCLWIGRQCDGTDNCGG GTDEEDCEPPTAHTFICKNBKEFLCRNQRCL SSSLRCNMFDDCGDGSDEECAFFV CPPRRPPRCKNBRVCLWIGRQCDGTDNCGG GTDEEDCEPPTAHTFICKNBKEFLCRNQRCL SSSLRCNMFDDCGDGSDEECAFFV CPPRRPPRCKNBRVCLWIGRQCDGTDNCGG GTDEEDCEPPTAHTFICKNBKEFLCRNQRCL SSSLRCNMFDDCGDGSDEECAFFV CPPRRPPRCKNBRVCLWGRDDPHA AVVDPLRGTMYWNDTGGSCASFGCTWR CLLSPGGVCCNNGCGCGCTAFFGGSGCANSFCT VNQGNQCCCLPGFTCBCRCGYCCASFGCCVNK CSRCLBGACVVNKQSGDVTCNCTDGRVAPS UTCRCGGCGGCGCAAFFGGSCANSFCT VNQGNQCCCLPGFTCBCCCCPCHMTGGCAASPGC MFTCCCGGCGGGGGCGCCTAFFGGSGCVNK CSRCLBGACVVNKQSGDVTCNCTDGRVAPS CLTCVGGCGGGGGGACAGGGGPCHMGLLLLLLULL VAGVVPWYKRRVCQAGAGGGCPTAMTGGGCAASPGC					sequence		
MSWTCDKEDDCEHGEDETHICNIFCESADACE CONHRCISKOWLOGSDEAGANCE GKTCGPSSTSCGTHVCVPERWLCDGNECA DGADESIAAGCL VINSTCDDEERNCONROCCIP KHPVCDHTDCADGSDESSPECEYPTCGFSEIF RCANGRCLSSRQWCCDGENDCHDQSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVABALLCN GQDDCGDSSDERGGHNECLSRGCODC EDLKIGFKCRCRPGFRLKDDGRTCADVDECS TTFPCSGRCTHTGSYKELCVEGYAPRGGDP HSCKAVTDEEPFLIFANRYVLKKLNDGSNY TLLKOGLNNAVALDFDYREQMIYVTDVTTTQ GSMERRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGRLYWCDKGRTTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNOYLVWTDWGDHSL IGRIGMOGSSRSVLVDTKTWPRGLTDVTTE RIYWADABEDVIEFASLDGSNRHVVLSQDIPH IFALTLEEDYVYWTDWTKISFNRSKTGTN KTLLISTLIRPRDLHVHALROPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTWWCDTF BCCCC THE DDCGDDCDDCDDNSDEANCHPWWCDTF DDCGDDSDEPPCFFCKCPGGCSTGICTN PAFICDGBNDCQDNSDEANCHPWWCDTF DDCGDDSDEPPCFFCKCPGGCSTGICTN PAFICDGBNDCQDNSDEANCHPWWCDTF DDCGDDSDEPPCFFCKCPGGCSTGICTN PAFICDGBNDCQDNSDEANCHPWCDDSDCVD GDEFANCTQRTCCOVERSCROSCCPARW KCDGEDCGDGSDEPKECPGCSTGICTON PAFICDGBNDCQDNSDEANCHPWCDDDVDVD GDFANCTQRTCCOVERSCROSCCPARW KCDGEDCGDSDEPCECCPGCCTGICTN PAFICDGBNDCQDNSDEANCHPWCDDDDVD GSDEFANCTQRTCCOVERSCROSCCPARW KCDGEDCGDGSDEPCECCPRCPGCCTGICTON PAFICDGBNCCCDDFRCCNGCCPARW KCDGEDCGDGGGGKGCCPARW KCDGBDCGDGGGGGGGCCPARW KCDGBDCGDGGGGGGGCCPARW KCDGBDCGDGGGGGGGCCPARW KCDGBDCGDGGGGGGGCCPARWCCD DADCMGGGGGGGGGCCPARW KCDGBDCGDGGGGGGGCCPARWCCD GGGGGGGGGCGGGCGGGCGGGGGGGGGGGGGGGGG							
CQNHRCISKOWLCDGSDDGGGSDEAAHC GKTCOPSSTSCTPIVCVPERWLCGDKDCA DGADESIAAGCLYNSTCDDREFMCQNRQCIP KHPVCDHRDCADAGSDESPECEPTGGPSEF RCANGRCI.SRQWECDGEADCHDQSDEAFY NPHCTSPEHKCHASSQFLCSSGGCVAEALLCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRPGFRLKDDGRTCADVDECS THTPCSSRCINTHGSYKCLCVEGVAPPIGGDP HSCKAVTDEEPFLIP ANRYTYLKIKLDGGSNP HSCKAVTDEEPFLIP ANRYTYLKIKLDGSNPY TILLRQGLNNAVALDFDYREQMITWTDVTTQ GSMIRRAMILNGSNVQVLHRTGLSNPDGLAV DWVGGRLYWCDKROTITEVSKLNGAYRTVL VSSGLREPRALVDDVQNGYILWTDWTDVTTQ GSMIRRAMILNGSNVQVLHRTGLSNPDGLAV ADWVGGRLYWCDSKROTITEVSKLNGAYRTVL VSSGLREPRALVDDVQNGYILWTDWTGSTRYNGLLDVYTE RIYWADAREDYIEFASLDGSNRIVVLSQDIPH IFALTLFEDYYVTYTDWTKISNRAMKITTGTN KTILLISTLIRPMDLHVFHALRQPDVPHHPECK VNNGGCSNCLLEPGGGHKCACFUNTYLGSD GRTCVSNCTASGPVCKNNKCIPFWWKCDTE DDCGGDHSDEPPDCPEFKCRGGCGCSTGICTN PAFICDGDNDCQDNSDEANCDHTVCLPSGK (CINTINRCINGIFRCNGQDNCCGDEDERDCPE VTCARNQPGCSTTKRCIPRYWVCDRDNDCVD GSDEPANCTGMTGOVDFFRCKSDGCPERARW KCDGEDDCGDDSDEPKEECDERTCEPYOFRR KCDGEDDCGDSDEPKEECDERTCEPYOFRR KCDGEDDCGDSDEPKEECDERTCEPYOFRR KCDGEDDCGDSDEPKEECDERTCEPYOFRR KCDGEDDCGDSDEPKEECDERTCEPYOFRR KCDGEDDCGDSDEPKEECDERTCEPYOFRR SCSSEFSCANGRCLAGRWKCDGDBIDCADD DECKTPCCMBOPGCKSGRCIPLAWRCDA DADCMGGSDEAACGTGTTCPLDEFQCNNI LCPLAWKCDGEDDCGNSDEESCTPR PCSSESFSCANGRCLAGRWKCDGDBIDCADD DECKTPRCMBOPGCKSGRCIPLAWRCDA DADCMGGSDEACGTGTTCPLDEFQCNNI LCPLAWKCDGEDDCGDSDEPEECARFV CPPNPFRCKNDRVCLWIGRGCDGTDNCGG GTDEEDCEPFTAHTHTCKNKERTLANGNCL SSSLRCMMFDCCGGSBEEDCSDPKLTSCAT NASICGDEACVATHTHCKNKERTLANGNCL SSSLRCMMFDCCGGSBEEDCSDPKLTSCAT NASICGDEACVATHTHCKNKERTLANGNCL SSSLRCMMFDCTGGSSVRIDAADVHYNAW WITDSGDVBEVAQMKGBNKELISAAMOGT LREIT.VONDNQWTTILNSGRLERSCAT NASICGDEACVATRICHLOWN CORRECT CPPNPFRCKNDRVCLWIGRCCGTDNCGG GTDEEDCEPFTAHTHCKNKERTLANGNCC SSCLEGGCVVNKQGGDVCCANNSTCT VNQGNQPQCRCLARGGRCTAATSPGRCC PRYTGDKCELDQCWCHRNGGTCAASPSS MPTCCRCPRGGGCCAANRSTCT VNQGNQPQCRCLARGRCTAATSPGRCCWA CSCLEGGCVVNKQGGGDVTCNCTGDGKCWAR CLLSSGGVVCCANNSTCT VNQGNQPQCRCLARGILGPLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL							
GKTCGPSSTSCGTHVCVPERWLCDGNEQCIP KHPVCDHDRDCADGSDESPECEYPTCGPSEF RCANGRCLSSRQWCDGENDCLDQSDEAPK NPHCTSPPHKCNASQFLCSSGRCVASALLC GODDCODSSDERGCHINECLSKKLSGCSQDC EDLKIGFKCRCRPGFHKDDGRTCADVDECS TTTPCSQRCINTHGSTYCLLCVEGYAPRGGDP HSCKAVTDESPELHARRYTYRKINLGGNDY TLLKQGLNNAVALDFDYERQMTYNTDVTTQ GSMRRMHLINGSNYQVLHETGLSRPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTYL VSSGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDVTTE PRIYWADAEDYIEFASLDGSNRHVVISQDIPH IFALTLFEDYYWTDWFIKSINRARKTITGTN KILLISTLHRRMDLHVFHALRQPDYPNHPCK VNNGGSNLCLLSPGGGHKCACTPTSYLGSD GRTCVSNCTASGFVCKNDKCIPPWWCDTE DDCGDHSDEPPDCPEFKCRPGGPGCSTIGLTN FALTOGDNDNCQDNSDEANCDHVCLPSQFK CTNTRCGFRCNGQDNCGGGEDERLOCE VTCARNQPCSITRTSCPFWWCDBDDCVD GSDEFANCTONTCGVDEFRCKDSGRCPAW KCDGEDGODGDSPFRECDBSTCEPYOFRC KNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEPSCANGRCAGRWKCDGDHDCDD GSDEFANCTOMTCGVDEFRCKDSGRCPAW KCDGEDGODGDSPFRECDBSTCEPYOFRC KNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEPSCANGRCAGRWKCDGBDEDCDG GTDEEDCGPTATITHTCKCKEFTLPRWCDA DADCMDGSDEAACGTGWTCPLDEFQCNN LCCPPNEPRCKNDRVCLWRGQCGTDNCGD GTDEEDCGPTATITHTCKCKEFTLPRWCDA DADCMDGSDEAACGTGWTCPLDEFQCNN CCPPNEPRCKNDRVCLWRGQCGTDNCGD GTDEEDCGPTATITHTCKCKEFTLPRWCDA NASICGGEARCWTTTHTCKCKEFTLPRWCDA NASICGGEARCWTTHTCHCKEFTLPRWCDA NASICGGEARCWTTTHTCKCKEFTLPRWCDA NASICGGEARCWTTTHTCKCKEFTLPRWCDA NASICGGEARCWTTTHCKCKEFTLPRWCDA CPPNEPRCKNDRVCLWRGQCGTDNCGD GTDEEDCGPTATITHTCKCKKEFTLANGCL SSSLRCMMFDDCGDSSVBLDAANDVTWKA GRYYMTWHTGTTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYMTWHTGTSYRSLPPAAPPTTSNYKA GRYYTONCGEDCCOPTORCCORCE PPATGDCCEPTCCWCRCGCGTCAASPS WCCCGRCCATAPFECGRCCYNCCCORCE PROTGCOMADDSGRCCATATEGGRCCYTACCORCE PPATGDCCEPTCCWCCGGCGCTTACCGRCCATAPFSCOC PPATGDCCEPTCCWCCGGCGCTTACFGGRCCYNK CSCCGGCGCGCTATCFGGRCCYNK CSCCCGCGCGCGCTATCTGCGCCCTATCGCGCCC PFCTCOWADDGGGGCCTTATCGGRCCTATCGCCC PFCTCOWADDGGGCCTTATCGGRCCTATCGCCC PPATGDCCC	ŀ		1				l
DGADESIAAGCL_YNSTCDDREFNCONROCIP  KHPVCDHRDRCDAGDSESPSECEPYTGGPSEP  RCANGRCI_SSRQWECDGENDCEDQSDEARP  MPHCTSPEHKCANSOQL_SSEGCVASALLCN  GQDDCGDSSDERGCHINECLSRKLSGCSQDC  EDLKIGFKCRCRPGFRLKDDGRTCADVDCS  TIFFCSQRCINTHOST SYKLCLVEGY APROGDP  HSCKAVTDEEPFLIFANRYTJRKLNLDGSNY  TLLKQGLNNAVALDFDYREQMITYTDWTDTY  GSMIRRMHINGSNVQVLHRTGLSNPDGLAV  DWVGGNL_YWCDKGRDTIEVSKLNGAYRTVL  VSSGLREPRALVVDVQNGYLJWTDWGDHSL  IGRIGMGSSSRSYVDTKITWPHOFLILDYVTE  RIYWADAREDYIEFASLDGSNR;VVLSQDIPH  IFALT_FEDVYVWTDWTKISFNRGLTLGTVYTE  RIYWADAREDYIEFASLDGSNR;VVLSQDIPH  IFALT_FEDVYWTDWTKISFNRGHKTGTIN  KTLLISTLHRMDLHVHHALRQPDYNHPECK  VNNGGCSNLCLLSPGGHKCACPTNYLGSD  GRTCVSNCTASQPVCKNDKCIPFWWCDTI  DDCGDDISDEPDGPEFKCRPGQCSTGICTN  PAFICDGDNDCQDNSDEANCDHVCLPQFK  CTNTRRCHGFRCNGQDNCGDGEDERDCPE  VTCARNQFQCSITKCFPWWCDDDCDVD  GSDEFANCTQMTCGVDEFRCKDSGCGFAK  KCDGEDDGODSDSPERECDERTCEPYQFRC  KNNRCVPGRWQCDVDNDCGDNSDESCTFP  PAFICDGDNDCQDNSDEPRECDERTCEPYQFRC  KNNRCVPGRWQCDVDNDCGDNSDESCTFP  POSESFSCANGRCGAGRWKCDGBDLCADGS  DEEDCTPRCDMDQFQCKSGHCPLBWRCDA  DADCMGSDEEACGTGWTKCDGBDLCADGS  DEEDCTPRCDMDQFQCKSGHCPLBWRCDA  LCPLAWKCDGEDDCGDNSDENFECCAFFV  CPNRPFRCKNDRVCLWIGRQCDBDCGDNSDESCTFP  POSESFSCANGRCGAGRWKCDGBDLCADGS  DEEDCTPRCLLGFQCNNI  LCPLAWKCDGEDDCGDNSDENFECCAFFV  CPNRPFRCKNDRVCLWIGRQCDFINCGD  GTDEEDCEPPTAHTTHCKDKEFLCRACT  ANSICGDEARCVYRTEKAAVCACRSGFHTVPG  QPCCQDNECLREFTCSQLCNTKGGHLCSC  SSLRCMPDDCGDGSDEEDCSDFKLTSCAT  NASICGDEARCVYRTEKAAVCACRSGFHTVPG  QPCCQDNECLREFTCSQLCNTKGGHLCSC  ARNFWKTHNTCKABGSEYQVLYIADDNEIRS  LPFGIPHSAYEQACPQODESVRIDAMDVIKVA  GRYWTNWHTGTISYRSLPAAPPTINNRIR  RQDRGVTHLNISGLKAPRGIADWNGHPKETAAMDGD  LEETL_VQDNQWFTCLAVPSTPDV  FEDYYGTYNNRYKHIKGGHLCSC  ARNFWKTHNTCKABGSEYQVLYIADDNEIRS  LPFGIPHSAYEQACPGOLGSVRIDAMDVIKVA  GLSSSQCCCACATGRCCATARGCPKCC  QPRYTORCCELDCQWGCGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACASSGRC  QPRYTORCCELOCOWERCHARGCACPHA  CUCCUGHCSNGSCTLANTSCMPRECQCPPH		i	1				
KHPVCDHIRDCADGSDESPECEYPTICGFSER RCANGRCLSSRQWCCGGENDCLDGSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVASALL, GODDCODSSEERGCHINECLSKLSGCSQDC EDLKIGFKCCRPGFILKDGRTCADVDEC. TITPCSQRCINTHOSYKLCLVEGYAPRGGDP HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY TLLKQGLNNAVALDFDYREQMTWTDVTYTQ GSMRRMHLMGSNVQLLHRGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRYTU VSSGLREPRALVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWFNGLTLDYVTY VSSGLREPRALVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWFNGLTLDYVTY TILLTLIFEDYYWTDWGTHSSNRARKTTGTN KILLISTLIHRMDLHVHALRQPDYPNHEPCK VNNGGSSNLCLLSPGGGHKCACTPTFYLGSS GRTCVSNCTASGPVCKNDKCIPPWWCDTE DDCGDHSDEPPDCPEFKCRPGGPGCSTGICTN KILLISTLIHRMDLONDSDEANCHINCLPSQFK CYNNTASGPVCKNDKCIPPWWCDTE DDCGDHSDEPPDCPEFKCRPGGPGCSTGICTN PAFICIODNOCQNDSDEANCHINCLPSQFK CTNTRRCHGFRCNGQDNCGDGEDERDCP VTCARNQPCSITRKCPRVWCDENDCVD GSDEFANCTQMTCGVDEFRCKDSGRCPAW KCDGEDGADGSDEPRECEDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCECTPP PCSESFSCANGRCAGRWKCDGEDEDCDCP VTCARNQPCGSTTRKCDFVWCDENDCVD GSDEFANCTQMTCGVDEFRCKDSGRCPAW KCDGEDGAGDGDEPRECEDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCECTP PCSESFSCANGRCAGRWKCDGBHDCADGS DEKDCTPRCMDQPQCKSGHCIPLRWCDA DADCMDGSDEEACGTGVTCDLEFCCNNT LCCPLAWKCDGEDDCCDDNSDESCECTP PCSESFSCANGRCAGRWKCDGBHDCADGS DEKDCTPRCMDQPQCKSGHCIPLRWCDA DADCMDGSDEEACGTGVTCDLEFCCNNT LCCPLAWKCDGEDDCCDDNSDESCECTP PCSESFSCANGRCAGWWKCDGBHDCADGS DEKDCTPRCNDQPQCKSGHCIPLRWCDA ADACMGSDEEACGTGVTCDLEFCCNNT LCCPLAWKCDGEDDCCDDNSDESCECTP PCSESFSCANGRCAGWWKCDGHDCADGS DEKDCTPRCNDRCAGRGCGTGVTCDNT LCCPLAWKCDGEDCCDDNTLTSCAT NASICGDEACWYTTHACKDKETCLRNORCL SSSLRCMMFDDCGDSSDEDCSDPKLTSCAT NASICGDEACWYTTKAGGHLCSCA GRYWTNWHITTSTYSLPPAPPTTSNRIR RQDGGVTHLNISGLKMPRGIADWWAGNNY WTDSGROTHLNISGLKMPRGIADWWAGNY WTDSGROTHLNISGLKMPRGIADWWAGNY WTDSGROTHLNISGLKMPRGIADWWAGNY WTDSGROTHLNISGLKMPRGIADWWAGNY WTDSGROTHLNISGLKMPRGIADWWAGNY WTDSGROTHLNISGLKMPGIADWYTNCTDGRKCCEWL CLLSSPCYCTCPNKRLDYGCACYACSSYC PRYTGMCTGLAVPTGTCQVPSPTPP PDAPRGTCNLQCPNGCGGCTAARRQPKCC QPRYTGMCTGLAVPTGTCTQWPSPTPP PDAPRGTCNLQCPNGCGGCTAARRQPKCC PRYTGMCTGWAGAGGGCTTAARRQPKCC PRYTGMCTCCPURCTTGRAWAB CLTCVGHCSNGGCGTTAARRQPKCCC PRYTGMCCCANNONDAGAGGCCTTATGTGSRCWA		l	'				I
RCANGRCLSSRQWECDGENDCHODSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN GODDCGDSSDERGCHINECLSKLLSGCSQDC EDIKJGFKCRCRGFGFLKDOGRTCADVDECS TTTPCSQRCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEFFLIRANRYTLRKINLDGGDP HSCKAVTDEEFFLIRANRYTLRKINLDGGDP HSCKAVTDEEFFLIRANRYTLRKINLDGGDP HSCKAVTDEEFFLIRANRYTLRKINLDGGDP HSCKAVTDEEFFLIRANRYTLRKINLDGAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSGRIEFRALVVDVQNGYLYWTDWGDAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGRIEFRALVVDVQNGYLYWTDWGDSL IGRIGMGGSSSSVIVDTICTWFRIGLTLDYVTE RIYWADAREDYIEFASLDGSNEHVVLSQDIPH IFALTLFEDYVYTDWETISSNRAHKTIGTIN KTLLISTLHEPMDLHVFHALRQPDVPHIPCK VNNGCSNLCLISFGGGHKCACPTNYTLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDIPE DDCGSHSDEPPOCEPFCKRPGGQCSTGICTN PAFICDGDNDCDDNSDEANCDHVCLPSQK CTNTNRCINGERCNGQDNCGEDEPDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDETANCTQMTCGYDEFRCKDSGRCIPAW KCDGEDDCDGSDEREDECDEEP VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDETANCTQMTCGYDEFRCKDSGRCIPAW KCDGEDDCDGSDEFRECEDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGBIDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMGSDSEAACGTGVRTCDEPQCNNT LCKPLAWKCDGEDDCGDNSDENFECARFV CPPNFFRCKNBRVCLWINGGCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNORCL SSSLECNMFDDCGJOSDEDNFECARFV CPPNFFRCKNBRVCLWINGGCDGTDNCGD GTDEEDCEPFTAHTHCKKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LPFGHFBAYEQAPQGDESVRIDAMDVIVAA GRYWTNWHTTISTRSLEFTAAPTTSNRFR RQDRGVTHLNISGLKMPRGLAUMONY WTDSGRDVEWAQMKGGENKTLISGMDEPH AIVVPLRGTHYYWSDWGNIPKETAMDGT LRETTLVQDNIQWPTGLAVDYNBRRLYWADA KLSVIGSRILNGTDFIVAADSKRGLSHPFSIDV FEDYTYGTYTJNNRVFKLHKFSPLVVLTGG LSHASDVVLYHOHKOPETNNCDKKCEWL CLLSPGGPCCTPGRRCLIDGTCVAPSPTDP DAPRRGTCNILQCPNGGSCFLNARRQRYCRC QPRYTGMKCHQWPTGLAVDYNBRRLYWADA KLSVIGSRILNGTDFIVAADSKRGLSHPFSIDV FEDYTYGTYTJNNRVFKLHKFSPLVVLTGG LSHASDVVLYHOHKOPETNNCDKKCEWL CLLSPGGPCCTPGTRCLIDGTCVAPSPSDV FEDYTYGTYTNNRVFKLHKFSPLVVLTGG LSHASDVVLYHOHKOPETNNCDRKCCWL CLLSPGGPCCTCPGRRCLIDGTCVAPSPGOC QPRTGDKCGTCTASPGGCCTHARRGCCAASPGG MPTCCCTGTGTTGPCCTTQQVCAGYCANNSTCT WQGONGPCCLCPGFRGTCLONGTCAASPGG CLTCVGHCSNGGSCTMNSKMMPECQCTPFIM TGPRCEEHVFSQQQFGHIASILIPLILLLLULLULV VAGVYWYKRRGGGCTMTAKMMPECQCTPFIM			İ				, , ,
NPHCTSPEHKCNASSOFLCSSGRCVAEALLCN GODDCGDSSBERGCHINECLSK.GGCODC EDILKIGFKCRCRFGFRLKDDCRTCADVDECS TITFPCSQCLINTHGSYKCLCVGYAPRGGDP HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY TLLKQGLNNAVALDFDYREQMYWTDVTQ GSMIRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGREPRALVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNYVLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNYCLSNGKCTFWWKCDTE BYTGLEDVYYWTDWETKSINRAHKTIGTN KTLLISTHEPRDLHYPHALROPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTHYTJGSD GRTCVSNCTASGYVCKNDKCDFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGPQCSTGICTN PARICDGDNDCQDNNSDEANCDHVCLPSGYK CTNTNRCHGEFRCNGQDNCGDGEDERDCPE VTCAPNGPGCSITKRCIPRVWVCDRONDCVD GSDEPANCTQMTCGVDEFRCKDSGRCPJARW KCDGEDDCGDSBPFKEECDERTCEPYGPC KNNRCVPGRWQCDYNDRCGDSDEESCTPR PCSESFSCANGRCIAGRWKCDGBDECADGS DEKDCTPRCDMOQPCKSGHCPJRAWCDA DADCMDGSDEAAGTGVRTCPLDEPCCNNT LCKPLAWKCDGEDDCGDNSDESCTPR CSESFSCANGRCIAGRWKCDGBDECADGS DEKDCTPRCDMOQPCKSGHCPJRAWCDA DADCMDGSDEAAGTGVRTCPLDEPCCNNT LCKPLAWKCDGEDDCGDNSDESCTPR CSESFSCANGRCIAGRWCCDGDDDCDNGC GTDEEDCEPT AHTTHCKDKKEFLCRNQCL SSSLECNHPDLOGGDSSDEPKECCDSTPRINTSCAT NASIGGDEARCVRTEKAAYCACRSGFHTVPG QPCQDNECLRGTGSGCCHNGGGLAGCN GRYWTWWTTGTISTYRSLPPAAPPTTSNRIR RQIDDGAGTBCGGSCFLNARRQCCC ARIYMKTHNTCKAGGECGVVLYJADDNERS LFFGHBASYRGAPGODESVNTRIRIR RQIDGRAVTHWHTGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPPTTSNRIR RQIDGRAVTHWATGTISTYRSLPPAAPTTSNRIR RQIDGRAVTHUATGTISTYRSLPPAAPTTSNRIR RQIDGRACGCCTATTFGGRCCHARRQCCCC PRYTGMCCCLCPGTLGTGCCAGCGGCCT			}				
GODDCGDSSDERGCHINECLSKLLSCSQDC EDILGIFKCRCRPFELLDORTCADVDECS TTFPCSQCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPFLIFANRYYLRKINLÖGEDP HSCKAVTDEEPFLIFANRYYLRKINLÖGEDP HSCKAVTDEEPFLIFANRYYLRKINLÖGEDP HSCKAVTDEEPFLIFANRYYLRKINLÖGEDP HSCKAVTDEEPFLIFANRYYLRKINLÖGEDP TLLKQGLNNAVALDFÜTREQMYTWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPFDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWEDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWEDGLAV SIGNEMOSSSSVIVDTKTITWFGLTLDYVTE RIYWADAREDYIEFASLDGSNEHVYLSQDIPH IFALTLEEDVYVYTDWETSISNRAHKTIGTIN KTLLISTHRPMDLHVFHALRQPDVPHFPCK VNNGGCSNLCLLSPGGHKCACPTNYFLGSD GRTCVSNCTASQFVCKNDKCLFFWWKCDTE DDCGDHSDEPPDCEPKCRPGGPCCSTGCTN KTLLISTHRPMDLHVFHALRQPDVPHFPCK VNNGGCSNLCLLSPGGHKCACPTNYFLGSD GRTCVSNCTASQFVCKNDKCLFFWWKCDTE DDCGDHSDEPPDCEPKCRPGGPCCSTGCTN FPARICGGNDCQDNSDEANDHVCLPSQFK CTNTINCIPGIFRCNGQDNCGDGEDEEDCEP VTCAPNQFQCSIKRCIPRVWCDRDNDCVD GSDEPANCTQMTGQDVEGPCBDCFDCF VTCAPNQFCCSIKTRCIPRVWCDRDNDCVD GSDEPANCTQMTGQDVDEPCCADGSCPAW KCDGEDDCGDPSDEPKEECRETCEPYQFRC KNNCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPRWRCDA DADCMDGSDEEACGTGVRTCPLDEPCONNI LCKPLAWKCDGEDDCGDNSDENEECARFV CPPNRPFRCNDRVCLWIGRQCDGTDNCOD GTDEEDCEPTHTHTHCKNKKEFLCRNORCL SSLACNMFDDCGDGSDEEDCSDDPKITSCAT NASICQDEARCCYTEKAAYCACRSGFHTVPG QPGCQDINECLRPGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQULYADADNEIRS LFPGHHSAYSGLAPPGAIDAWAGNVY WTDSGRDVEVAQMKGENRKTLISGMDPPH AIVVPLRGTNYWSDWGNIPSGLIPPFSDN PGDYTYGVTYNNWTGTISTYRSLPPAPPTTSNRIR RQDRGVTHLNSGLKMPRGJADWAGNVY WTDSGRDVEVAQMKGENRKTLISGMDPPH AIVVPLRGTNYWSDWGNIPSGLIPPFSDV PDAPRFGTCNLQCNGGSCTLNARRQFCCC QPRYTGDKCELDQCWEHCRGGTCAASPSG MPTCRCTOPIGRALDNGTCVPVSPTPP PDAPRFGTCNLQCNGGSCTLNARRGKCCWL CLLSSEGVCLTCPGGSCCTLNARRGKCCWL CLLSSEGVCLTCPGGSCCTLNARRGCCCACASPSG MPTCRCTTOFTGFKCTQQVCAGGCANNSTCT VNQGNQPOCCRLPGGLTGACASPSG MPTCRCTTOFTGFKCTQQVCAGGCANNSTCT VNQGNQPOCCRLPGGLGGCCTNATSGRACPNAC QPRYTGDKCELPGGLGGCCTNATSGRACPNAC CLTCVGHCSNGGSCTMNSKMMPECQCPPFM TGPRCEEHVFSQQQFGHLASILIPLLLLLLLULLULL VAGVYFWYKRRVQCAAKGFGRMTNOAM	İ						
EDILIGIFICACORTOGENTA DIVIDECS TTPECSQRCINTINGSYKCL VOYAPRGGDP HSCK AVTDEEPFLIRANRY YLRKLINLDGSNY TLLK QCLNNAV ALD POYAPRGGDP GSMRRMHLNOSNY QVLHRTGLSNPDGLAV DWVGGNLY WCDKGRDTIEV SILNGAYRY VL VSGGLREPAL VVDVQNGYLLY WTDWGDHSL IGRIGMDGSSRS VIVDTKITWPHGLILDY VTE RIYWADARED YIPE ASLDOSNRHVVLSQDIPH IFALTLEED VYVYWTD WETKSINRAHKTIGTN KTLLISTHEPMDLHYPHALROPDYPHHPCK VNNGGCSNLCLLSPGGGHKCACPTNYTLGSD GRITCYSNCTASGYVCKNDKCDTE DDCGDHSDEPPDCEPFKCRPGOPQCSTGICTN PAFICDGDNDCQDNSDEANCDHYCLPSCK VNNGGCSNLCLLSPGGGHKCACPTNYTLGSD GRITCYSNCTASGYVCKNDKCDTE DDCGDHSDEPPDCEPFKCRPGOPQCSTGICTN PAFICDGDNDCQDNSDEANCDHYCLPSCK CTNTNRCHGIFRCNGQDNCGDGEDERDCPE VTCAPNQPGCSITKRCIPRVWCDRONDCVD GSDEPANCTQMTCGYDEFRCKDSGRCIPARW KCDGEDDCGDSBEPKEECDERTCEPYOFRC KNNRCYPGRWQCDYDNDCGDSBEESCTPR PCSESETSCANGRCIAGRWKCDGDBDCADGS DEKDCTPRCDMDOPQCKSGHCPLRWCDA DADCMDGSDEEAGGTGVRTCPLDEPGCNNT LCKPLAWKCDGEDDCGDNSDENPECARPY CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTHCKNKKEFLCRNCRL SSILRCNMPDDCGDGSDEEDSUPRLISCATT LCKPLAWKCDGEDDCGDNSDENPECARPY CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTHCKNKKEFLCRNCRL SSILRCNMPDDCGGSSEEDSUPRLISCATT NASICGDEARCVRTEKAAVCACRSGFHTVPG QPCCQDNECLREGTCSQLANDMYHKKA GRYYWTNWHTGTIS YRSLPPAAPPTISNRIR RQIDGSCHLRIFT GRANDWYHKA GRYYWTNWHTGTIS YRSLPPAAPPTISNRIR RQIDGSCHLNGRAGMEPH ALVVDPLRGTMYWSDWGNEPKEITAAMDGT LRETLVQDNIQWPTGLAVYHNERLYWADA KLSVIGSRLNGTDPTVAADSKRGSHPPSDV FEDYIYGYTYNNWFTGHTSGRHCCS QPRYTGNCELTPFLORTCVPSPSTDV FEDYIYGYTYNNWFTGHTSGRHCPSC QPRYTGNCELTPFLORTCVPSPSTDV FEDYIYGYTYNNWFTGHTSGRHCPSC QPRYTGNCELTPFLORTCVPCSPSTDV FEDYIYGTYTNNWFTCHTSGRHCPSC QPRYTGNCELTPFLORTCRCGCCASGCC QPRYTGNCELTPFLORTCRCCYPGCSC NPGTCQMAADGSRQCCCTAYFEGSRCCVNK CSRCLEGACVVNKQSGVTCNCTGCRAPS QPRYTGNCELTPFLORTCRCCAPC QPRYTGNCELTPFLORGCCANAPSC CLTCVGHCSNGGSCTMARRQPKCRC QPRYTGNCELTPFLORGCCANAPSC CLTCVGHCSNGGSCTMASKMPECQCPFHM TGRCEEHVFSQQQPGHASLILPLLLLLLLLLLLL VAGWVFWYKRRQCAAGGGVRCNANTOCT VNQGNGPOCCACPTGLAGRADMATOCHT TOPRCEEHVFSQQQPGHASLILPLLLLLLLLLLLLLLLL VAGWVFWYKRRQCAAGGGGCTTAYFEGSRCCVNK CSRCLEGACVVNKQSGGVTCNTGRANTCT	1						
TTFPCSQRCINTHGSYECLCVEGYARRGGDP HSCKAVTDEEPFLIFANFLYRIKINLDGSNY TLLKQGLNNAVALDFDYREQMTWTDYTTQ GSMTRRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNLYWCDKGRDTEVSKLNGAYRTV1L VSSGLREPRALVVDQNGYLYWTDWGDHSI. IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVNLSQDIPH IFALTLFEDYVWTDWETSINNAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSALCLLSPSGGHCACPTNYTLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPCQPQCSTGICTN PAFICOGDNDCQDNSDEARCOCHOCHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEBDCDCFE VTCAPNQFQCSITRKCIPFWWCDRDNDCVD GSDEPANCTOMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDOSDEPRECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDERSCIPARW KCDGEDDCGDOSDEPRECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDENSECHARW KCDGEDDCGDSSDENERCCHRCCDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEPCCNNT LCKPLAWKCDGEDDCGDNSDENNEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNGGD GTDEEDCEPTAHTTHCNKEFFLCNNGRCL SSSLRCNMPDDCGGGSDSDENERCARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNGGD GTDEEDCEPTAHTTHCNKKFFLCRNGRCL SSSLRCNMFDDCGGGSDEDCSDPKLTSCAT NASICGDEARCVNTEKAACCRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNPMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPISAYEQAFQGDESVRDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTISNRIR RQIDRGYTHINSGLKMPRGIADWVAGNNY WTDSGRDVIEVAQMKGENRICLISGMIDEPH AIVVDPLRGTMYMSDGWFREETAAMDGT LRETLVQDNQWFTGLAVDYHRERLYWADA KLSVGSRIKGNTDPIVAGKGGLSCFTNARRQFKCRC QPRYTGMTWSDWGRKEGISHPFSDDV FEDYIYGVTYNNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQIKQPCFVTNPCGRKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPPSSTPP PDARPRGTCNLCCPNGGSCFLNARRQFKCRC QPRYTGMKCELDQCWEHRGGTCAASPSG MFTCRCTGFTGFKCIQQCAGYCRQCSGYCE NFGTOQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGIGNGGSCTNNSKMMPECQCPPHM TGPRCEEHVSQQQPGHAISLILLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL	İ						l '
HSCKAVTDEPFLIFANRYYLRKINLDGSNY TLLKQCI,NNA VALDFDYGMYWTDVTTQ GSMIRRMHLNGSNYQVLHRTGLSNPPGILAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWRNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYYYWTDWETKSINRAHKTGTN KTLLISTLHRPMDLBYFHALRQPDYPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKQFFWWKCDTE DDCGDHSDEPPDCPEFKCRPQGPGCSTGICTN PAFICDGDNDCQDNSDEANCDHVCJPSQFK CTNTNRCIBFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCCDRDNDCVD GSDEPANCTGMTGGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYOFRC KNNRCVPGRWQCDYDNDCDNDDEDSEECTPR PCSSESFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQPQCKSGHCIPLRWRCDA DADCMDGSDEPACHGVTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPPRCKNDRVCLWGRQCDGTDNCOD GTDEDCEPPTAHTTHCXDKKEFLCRANGRCL SSSLRCNMFDDCGDGSDEDECSIDPKTISCAT NASICGDEARCVRTEKAAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNYMKTHNTCKAEGSEYQLYJYADDNERS LF7GHPISAYEQAFQGDESVRIDAMDHVKA GRYYWTNWHTGTIST SREPPAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADDWYGNA GRYYWTNWHTGTIST SREPPAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADDWYGNA KISVIGSRIKHTCKAEGSESYQLVYHADDNERS LF7GHPISAYEQAFQGDESVRIDAMDHVKA GRYYWTNWHTGTIST SREPPAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADDWYGNA GRYYWTNWHTGTIST SREPPAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADDWYGNA GRYYWTNWHTGTIST SREPPAPPTTDRIRR RQIDRGVTHLNISGLKMPRGIADDWYGNA KISVIGSRIKGTHPYDAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADDWYGNA KISVIGSRIKGTMYFYGKARDGTCANSPG MPTCCFTGFTGFRCTQQVCAGYCANNSTCT VNQGNQPCCLPGFLGBRCQVRQCGYCANNSTCT VNQGNQPCCRLOFGLGBRCQVRQCGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNICOSGDYTCOTTCDRCVABS CLTCVGHCSNGGSCTNNSKMMPECQCPPHM TGFRCEEHVFSQQPGHHASILLILLLLLLL VAGVVYWYKRRVQGARGCRCTATYFEGSRCEVNK CSRCLEGACVVNICOSGDYTCOTTCDRCVABS CLTCVGHCSNGGSCTNNSKMMPECQCPPHM TGFRCEEHVFSQQPGHHASILLILLLLLLLLLLLL VAGVVYWYKRRVQGARGCRCTGATFGCGRCTANSCANSC							
TLIKQGLNNA VALDFDYREQMYUTDYTTQ GSMRRAHLMOSNY OVLAFIGLSNPDGLAV DWVGGNLY WCDKGRDTIEVSKLNGAYRTVIL VSGALREPRAL VVDVQNOTLYWTDWGDHSL IGRIGMGDSSS SNYIDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVYLSQDIPH IFALTLFEDYYWTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDYNHPCK VNNGGCSILCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQPVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGOFQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCIEGEFRCNGQDNCGDGEBRDCPE VTCAPNOFQCSITRACIPRWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCPARW KCDGEDDCGDSDEPKECPPGFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPARW KCDGEDDCGDSDEPKECRDFYGFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPARW KCDGEDDCGDSDEPKECRDGHDCADGS DEKDCTPRCDMDQPQCKSGFICPLRWRCDA DADCMDGSDEEACGTGVTRCDHDCADGS DEKDCTPRCDMDQPQCKSGFICPLRWRCDA DADCMDGSDEEACGTGVTRCDHDCADGS GTDEEDCEPTAHTTHCKDKEFI-CRNQRCL SSSLRCNMPDDCGDGSDEDCGDNSDENPECARRV CPPNRFFRCKNDRVCL WIGRQCDGTDNGGD GTDEEDCEPTAHTTHCKDKKEFI-CRNQRCL SSSLRCNMPDDCGDGSBDEDCSIDPKLTSCAT NASICGDEACVATEKAACCRSGFHVYPG QPGCQDINECLRPGTCSQLCNNTKGGHLCSC ARNFWKTINTCKAEGSSYQU-YLADDNERS LFPGHPBAYEQAPQODESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTINRHR RQIDRGVTHINSGLKMPRGJADDWAGNNY WTDSGRDVEVAQMKGENRTLISMDEPH AIVVDPLRGTMYMSDWGNEPKEITAAMDGT LREILVQDNQWFTGLAVDYHNERLYWADA KLSVIGSRIGNTOPIVAAGNGFISFISDDV FEDYIYGVTYNNRVFKHRFGGISPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCEWL CLLSPSGPVCTCPHGKRCLDNCTCVPYPSPTPP PDARRPGTCNLCCPNGGSCFLNARRQPKCRC QPRYTGNKCELDQCWEHROGGTCAASPSG MFTCRCTFGTFGFKCTQQVCAGYCANNSTCT VQGNQPQCRCLPGFLGBRCQVRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNASSPSG MFTCRCTFGTFGFKCTQQVCAGYCANSTCT VQGNQPQCRCLPGFLGBRCQVRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNASSRCEVHK CSRCLEGACVVNKQSGDVTCNASSRCEVHK CSRCLEGACVVNKQSGDVTCNASSPG MFTCRCTFGTFGFKCTQQVCAGYCGCGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVHK CSRCLEGACVVNKQSGDVTCNTGGACH			<u> </u>				
GSMRRNHLINGSNYQVLHRTGLSNPGLAV DWVGGNLYWCDKGRTJEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWFNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLEPDYYWTDWETKSINRAHKTGTN KTLLISTLHRPDLHYHALRQPDYPNHPCK VNNGGCSNLCLLSPGGHKCACPTHYLGSD GRTCVSNCTASQPVCKNBCPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCJPSQFK CTNTNRCIPGFRCNGQDNGDGDGEDERDCPE VTCAPNQFQCSITKRCPPRWWCDDRDDCVD GSDEPANCTOMTGGVDEFFRCNSGRCIPARW KCDGGDDCGDGSDEPKEECDERTCEPYQFRC KNNRCYPGRWQCDYTDNGCDDSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTFRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACCTGVCFLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENEECARFV CPPNRPFRCKNDRVCLWGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFI.CRNQRCL SSSLRCNMFDDCGDGSDEEDCSDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNYMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPBSAYEQAFQGDESVEIDAMUHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRIR RQIDRGVTHLNISGI.KMPRGIADWVARGNV WTDSGRDVEVAQMKGENKTLISGMIDEPH AIVVDPLRGTMYWSDWGNEPKEITAAMDGT LRETLVQDNQWFTGLAVDYHNERIYADAA KLISVIGSRINGTDFIVADSKRGS.SIFFSIDV FEDYTYGVTYNNRVFKHKFGHSFJVVLTGG LSHASDVVLYHQHKQPEVTNPCDRKCECWL CLLSPSGPVCTCPNGKRLDNGTCVPVSPFTPP PDARRFGTCNLQCFNGGSCFLNARROPKCRC QPRYTGGKCLDQCWEHCRNGGTCAASPSG MPTCCCTTGGTGFRCCQCAGYCACSGYCE NFGTCQMAADGSRQCRCTAYFEGRCPAAPS CLTCVGHCSNGGSCTNNSKMMPECQCPPHM TGFRCEEHVFSQQQPGHHASILILLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL		ļ	j				
DWVGGNLYWCDKGRDTIEVSKINGAYRTVL VSSGLREPRAL VVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRIVVLSQDIPH IFALTLEPDYYWTUDWETKSNRAHKITGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGGSNLCLLSPGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCPPFWKCDTE DDCGDHSDEPPDCPEFKCRPQQFQCSTGICTN PAFICOGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCPGIFRCNGQDNCGDEDERDCPE VTCAPNQCQSITKRCIPVWVCDRDNDCVD GSDEPANCTQMTGQVDETRCKDGGEDERDCPE VTCAPNQCQSITKRCIPVWVCDRDNDCVD GSDEPANCTQMTGQVDETRCKDGGEDERDCPE VTCAPNQCQSITKRCIPVWVCDRDNDCVD GSDEPANCTQMTGQVDETRCKDGGEDERDCPE VTCAPNQCQSITKRCIPVWVCDRDNDCVD GSDEPANCTQMTGQVDETRCKDGGEDERDCPE VTCAPNQCQSITKRCIPVWVCDRDNDCVD GSDEPANCTQMTGQCDGNSDEDESCTPR PCSESEFSCANGRIAGRWKCDGHDCADGS DEKDCTFRCMDDGPCKSGHCPLRWRCDA DADCMDGSDEEAGGTGVTTCPJDEFQCNNT LCKPLAWCDGEDDCGDNSDENFEECARFV CPPNRPFRCNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFILCRNQRCL SSSLRCMFDDCGDGSDEEDCSIDPKLISCART NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDNECLRETGCJCNNTKGGHLCSC ARNFMKTHNTCKAEGSSEYQVLYADDNERS LFPGHPISAYEQAPQGDESVRIDAMDVHVKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRIR RQIDRGVYHLNISGLKMPRGIADWVAGNVY WTDSGGROVENARGHGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNHPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPTVADSKRCISHPSIDV FEDYYGVTYNNRVYKHHRFGHSLVNLTGG USHASDVVLYHQHKQFEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSFTPP PDAPRPGTCNLQCPGGSCFLNARRQFKCCC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTTGTFGKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPFGLGRCCYNACGSGCC NFGTCQMAADGSRQCRCTAYFGSRCEVNK CSRCLEGACVNNKGSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQPGHLASILPLLLLLLLLL VAGVVFWKRRVQGAKGFQHCMTNGAM			١.		:		
GRIGMDGSSRSVIJDTKITWPNGLTLDYVTE RIYWADAREDYIER SLDGSNRIVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTIGTN KTLLISTLIRPMDLHVFHALRQPDVPNHPCK VNNGGGSNLCLLSRGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNIKCIPFWKCDTE DDCGDHSDEPPDCFEFKCRFGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIFGIFRCNGQDNCGJGEDERDCPE VTCAPNQFQCSITKRCFRVWVCDRDNDCVD GSDEFANCTQMTCGYDEFRCKDGRCPARW KCDGEDDCDGDSDEFKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTFR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTFRCMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRGTGCLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYADDNERS LFPGHPHSAYEQAFQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNHPKETAAMDGT LRETLVQDNIQWPTGAWYSDWGNHPKETAAMDGT LRETLVQDNIQWPTGAWYSDWGNHPKETAAMDGT LRETLVQDNIQWPTGVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPYPSTTPP PDAPRPGTCNLQCFRGGSCTLNARQFKCCEWL CLLSPSGPVCTCPNGKRLDNGTCVPYPSTTPP PDAPRPGTCNLQCFRGGSCTLNARQFKCCC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTTGTFTGFRCTQQVCAGYCANNSTCT VNQGNQPQCRCLPFGLGRCCYNQCSGYCE NFGTCQMAADGSRQCRCTAYFGSRCEVNK CSRCLEGGCVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQFGHLASILLLLLLLLLLLLV							
RIYWADAREDYIEFASLOGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNYTLGSD GRTCVSNCTASGPVCKDMCCIPFWKCDTE DDCGDHSDEPPDCPEFKCRPGGPCCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITRCIPRVWCDRDNDCVD GSDEPANCTQMTCOPFERCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGHDLADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA ADADCMGSBEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRCDGTDNCGD GTDEEDCEPPTAHTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDSIDPKLTSCAT NASICGDEARCVATEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQODESVRIDAMDVHVKA GRYYWTNWHTGTISYRSLPAPPTTSNHR RQDREGVTHLNSGLKMPRGIADDWAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTIMYWSDWGNIPKETAAMDCT LRETLYQDNIQWPTGLAVDYNNERLYWADA KLSVIGSIRLINGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPSVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPARFGICNLCCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLAPTAFGSSRCEVNK CSRCLEGACVVNKQGSGVTCCNTDGRVAPS CLTCVGHCSNGGSCTLNASKMMPECOCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQRQMTNGAM	Ì					•	VSSGLREPRALVVDVQNGYLYWTDWGDHSL
IFALTLEDYYYWTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNIFPCK VNNGGCSNLCLLSPGGHKCACTNTYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDEFKCRPQGPCSCTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNQQDNCGDGEDERDCPE VTCAPNQFQCSTIKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCCPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGHDLCADGS DEKDCTPRCDMDQCKSGHCIPLRWRCDA DADCMDGSDEBACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKNKEFLCRNQRCL SSSLRCNMPDDCGGSDEEDCSDPFLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPQ QPCCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLTADDMEIRS LFPGIPHSAYEQAFQCDESVRIDAMDVHVKA GRYYWTNHTGTISYRSLPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWDGMSPREITAAMDGT LREILVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGOTYJNNRYKHIKFGHSPLVNLTGG LSHASDVVLYHQHK.QPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWHERCRGGTCAASPSG MPTCRCPTGFTFGFKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGGACVVNKQGSDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQFGHAISILIPLLLLLLLLL VAOVVFWYKRRVQGAKGFQRRMTNIGAM			İ		_		
KTILISTLHRPMDLHVFHALRQPDVPNHFCK VNNGGCSNLCLLSPGGHKCACPTNY*LGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICIGGNDCQDNSDEANCDHIVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCTPARW KCDGEDDCGDDSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR KCDGEDDCGDDSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDGFQCKSGHCIPLRWRCDA DADCMDGSDEBACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFPRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDSIDPKLTSCAT NASICGDEARCVATEKAAVCACRSGFHTVPG QPGCQDDNECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEQAFQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAPPTTSNRHR RQDRGVTHLNSGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPSIDV FEDYIYGTYTINNTSYKHKRFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLIDNGTCVPVPSPTPP PDAPRFGTGNLCCPFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFGCCCAPROSCSYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGGACVVNKSQGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTLNASKMMPECOCPPHM TGPRCEEHVFSQQQPGHLASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHRQMTNGAM	1 .		l		_		
VNNGGCSNLCLLSPGGHIKCÁCPTNYTLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCIPGIFFCKOQDNCGGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDEPTCPFYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTTRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGDDCGDNSDEPECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKKKEFLCRNQRCL SSSLRCNMFDDCGDSSDEEDCSINPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNYMKTINTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGITSYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNHPKETAAMDGT LRETLVQDNIQWPTCLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGISHPFSIDV FEDYITGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCCPTGFTGPKCTQVCVPVPSPTPP PDAPRFGTCNLQCNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCCPTGFTGPKCTQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSREVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHLASILIPLLLLLLLLV VAGVVFWYKRRVQGAKGPQQRGTNGAM		1		ļ	ļ		IFALTLFEDYVYWTDWETKSINRAHKTTGTN
GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGWRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPTAHTTHCKDKKEPLCNRCL SSSLRCNMFDDCGDGSDEEDCSIPFKLTSCAT NASICGDEARCVTREKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTENTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYVWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIADDWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTFGTFGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGBRCQYRCCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGPQQRQTNGAM	1						
DDCGDHSDEPPDCPEFKCRPGQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCFARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSSEFSCANGRCIAGRWKCDGHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTERAAYCACRSGFHTVPG QPGCQDNECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPHSAYEQAFQODESVRIDAMDVHVKA GRVYWTNWHTGTISVRSLPPAPPTTSNHR RQIDRGVTHLNISGLKMPRGIADWVAGNVY WTDSGRVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRRNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYNNRVFKHIKFGHSPLVNLTGG LSHASDVVLYHQHKQFEVTNPCDRKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWHCRNGGTCAASPSG MPTCRCTTGTTGVCCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQVRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTGRGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLLVL VAGVVFWYKRRVQGAKGFQHRMNTNGAM	į			į			
PAFICOGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNOPCQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGR WQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDERECARFV CPPNAPPRCNDRVCLWIGSQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSSDEEDCSIDPKLTSCAT NASICGDEARCVTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARIPMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPISASYCAFQGDESVIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRVUEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPPSIDV FEDYIYGVTYINNRVKHIKFGHSPLVNLTGG LSHASDVVLYHQIKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVSPPTPP PDAPRPGTCNLQCFNGGSCTLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTTGTTGFTCPKCTQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPEQCCPPHM TGPRCEENVFSQQQPGHASILIPLLLLLLLLLL VAGVVFWYKRRVQGAKGFQHROMTNGAM	ļ						
CTNTINCIPGIFECNGODNCGDGEDERDCPE VTCAPNOFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAGSESYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIADWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIEPKIETAAMDGT LRETLVQDNIQWFTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDVTYGVTYNNRVFKHIKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTPDCDKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARPRGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTIMNSKMMPEQCCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLLLLL VAGVVFWYKRRVQGAKGFQHRMTNGAM	1						
VTCAPNOFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPAR W KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCYPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKCTPRFCMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECAFFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNORCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTTSYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRINGTDPIVAADSKRGLSHFFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSSEGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCFTGFTGFKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGBRCCYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLL VAGVVFWYKRRVQQAKGFQRICNMTNGAM	i		ļ	!			
GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFI-CRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAVCACRSGFHTVPG QPGCQDDELRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCCPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLVL VAGVVFWYKRRVQQARGFILLLLLLLLLVL	İ		Ì				
KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFPRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIEPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSIPFSIDV FEDYTYGTYTNNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCKCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEHVPFSQQQPGHASILIPLLLLLLVL				ļ			
KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFPCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEGAFQGDESVRIDAMDVHVKA GRYYWTNWHTGTISVRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINDRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLVL	]						
PCSESEFSCANGRCIAGRWKCDGBHDCADGS DEKDCTPRCDMDQFGCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVI,YIADDNEIRS LFPGHPHSAYEQAFQODESVRIDAMDVHVKA GRVYWTNWHTGTTSYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYITGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVFVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGBRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGSSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHLASILPLLLLLLVL VAGVVFWYKRRVQQAKGFQHRMTNGAM	1		ł	i			1
DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEDCSIDPKLTSCAT NASICOGEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VQQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGSCTMNSKMMPEQCCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLLV VAGVVFWYKRRVQGAKGFQHQRMTNGAM	İ	1	l	ĺ	}		
DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPFTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKLETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANSTTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHASILIPLLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCFTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILPLLLLLLLLLV VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ļ						<b>1</b>
CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCFTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILPLLLLLLLLLV VAGVVFWYKRRVQGAKGFQHQRMTNGAM							LCKPLAWKCDGEDDCGDNSDENPEECARFV
SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				i			
NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPITSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	İ		i				GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPITSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKLHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM					ļ		
ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			1	[	į		
LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				I			
GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			İ				LIPGHPHSAYEQAFQGDESVKIDAMDVHVKA
WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM		ĺ	ĺ	1	[	(	
AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ŀ					1	
LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				}	!		
KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				1			
FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1	}	]	ļ	j		
LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	[			1			
CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1					l	
PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				1			
QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	}	j	}	]	1		
VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM		1		1			
NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1		1		}		VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ľ					•	CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
VAGVVFWYKRRVQGAKGFQHQRMTNGAM		}	1	1	ł		
				1			
NVEIGNPTYKMYEGGEPDDVGGLLDADFAL					1		
		<u> </u>	<u> </u>	<u> </u>			NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DPDKPTNFTNPVYATLYMGGHGSRHSLASTD
374	1724	A	3187	191	1815	EKRELLGRGPEDEIGDPLA  CLELASAGKIPEESKALSLLAPAPTMTSLMPG AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL NSQTTTADQLLEFFKQVGEVKFVRMAGDET QPTRFAFVEFADQNSVPRALAFNGVMFGDRP LKINHSNNAIVKPPEMTPQAAAKELEEVMKR VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK SHSRRKRSQSKHRSRSHNRSRSQKDRRRSK SPHKKRSKSRERKSRSRSHSRDKRKDTREKI KEKERVKEKDREKEREREKEREKERGKN KDRDKEREKDREKDKEKDREREREKEHEKD RDKEKEKEQDKEKEREKDRSKEIDEKRKKDK KSRTPPRSYNASRRSRSSSRERRRRSRSSSRS PRTSKTIKRKSSRSPSPRSRNKKDKKEKERD HISERRERERSTSMRKSSNDRDGKEKLEKNST S
375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP WF\FFYP*SPDLQIPSSFRNGLNDWYHHSQKC PDLDGAYVKK
376	1726	A	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K
377	1727	A	3201	274	1285	KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY
378	1728	A	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS
379	1729	A	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK *LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						XKQSNNKGTRASSYXEPDAXEQWKFPHKKL QLPGXTHE
380	1730	A	3207	187	507	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP AXLLPGPGGGPGPVASLEARAQASSGVTPNG GGRTYPYPTFSSGE
381	1731	A	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/ EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI KLQDMEKKANPSSLVLERREVEQQGFLHLGE HDGSLDLRSRRSVQEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHHHQGHNS LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF LFSNACL
383	1733	A	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK
						QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI
384	1734	A	3242	3	678	IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI
385	1735		3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRILLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  MTCADPGEIANGHRTASDAGFPVGSHVQYRC
						LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF SNPLYEAGDTREYEVSI
386	1736	Α	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGFFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMO
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDONKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQ\EELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN KQQQ  RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
						WICLSMVILTHSLKTFHRNWDWESEYTLFMS ALKVNKNNAKLWNNVGHALENEKNFERAL KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL WSEACAFL*AAAPQGPASPCCGLPSGFPRVW AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY STSFLTDSYLKYIGWTLHDKHREVRVKCVKA LKGLYGNRDLTARLELFTGRFKDWMVSMIV DREYSVAVEAVRLLILILKNMEGVLMDVDCE SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL TGAALAGSYPIWENENTLSWLPTFTYNFCLST PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI NILPPNQTILISVEASISSSPIRNKWALHLITLLT GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE DMHTSITSLQRQLDFLVGVILQNWRVLDLLT TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH DRAAEL*HQVADSWWQGSSLLRWIPWVAPF LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRRELEIATSDNQE YYNRLCQEVTNRERNDQKMLADLDDLNRTK KYLEERLIELLRDKDALWQKSDALEFQQKLS AEERWLGDTEANHCLDCKREFSWMVRRHHC RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP EKIVLRALKDSRAGMPEQDKDPGVQENPDD QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF VVRTMCAVLGLVARQEDSGLRDHSVRVLISN HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		P.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	l	l	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	<del> </del>	<del> </del> -	<del> </del>	sequence		nucleotide insertion
		l				GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ
•		İ				FFFFVIFLNIFLLAFSSPGSQPLLNSPPSFVCWSR
						GFMEMNGRGELVESLKRFCASTRLPPTPLLLF
j	]	}				PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ
					1	VQRTLVSVTVSDASWVSELL\WSLFVPFTVY
					Į.	QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE
1		ł			}	LG\QTGTRLTPA\DKAEHMKRQRHPR\LRPQS
						AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF CPHVAIGVFIPERPWPKTGCCKTLTIHLILL+G
1			1			GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG
!	1		1		ļ	PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL
· ·	İ	Í				QERKQ\ALYEYARRFTERRAPGGLD
403	1753	A	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS
		J	, , ,	1	'7'	GGASAGLASSPECACGRSHFTCAVSALGECT
						CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL
1						DFHCDNGKCIRRSWVCDSDNDCEDDSDEOD
						CPPRECEED
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG
		ŀ			ĺ	QDHVQNEEIYARVLDKFGSNFLSRDNADLGT
						AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS
					ļ	LLKGDLKGVKGDLKKPFDKAWKDYETKFAK
1_	1	1	[		ĺ	IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA
İ						KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG
ļ			]			HPASKENDQMVDTIKNTTKVPIIWTYGDMVE
		1				PRPQMIRPAVGAKHKELWKILMALKKIK\IWE
	<u> </u>					GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV
l		1				MCVLLWALSLLQSILEWMFCSFLFSDVDSDN
						WCQILDFLTAVWLIFLI\LVLCGFTLVLLVRIIC
]	ļ	1				GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F
407	1757	A	3328	213	1841	LLYWIEKDLDDL
107	1757	^	3320	213	1041	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY
		l				RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC
ľ	Ì	ľ	Ì			PHISFTGNICVYCPGGPDSDFEYSTOSYTGYEP
•						TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD
						KVEFIVMGGTFMALPEEYRDYFIRNLHDALS
						GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC
	[					MKRHLSDMLTYGCTRLEIGVQSVYEDVARD
						TNRGHTVKAVCESFHLAKDSGFKVVAHMMP
}	1					DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP
	J					TLVIRGTGLYELWKSGRYKSYSPSDLVELVA
	}					RILALVPPWTRVYRVQRDIPMPLVSSGVEHG
	1					NLRELALARMKDLGIQCRDVRTREVGIQEIH
	1					HKVRPYQVELVRRDYVANGGWETFLSYEDP
						DQDILIGLLRLRKCSEETFRFELGGGVSIVREL
						HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA
	·				į	ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL
	L					QGPYMVKMLK
408	1758	Α	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK
						KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG
						LKGRVFEESLADLQND\TDGYLLRVI*VAFTT
						ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH
400	1050				- 3	DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL
						WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK
				į		KGIRARILETLGMLLLLALLILGIVWVASALID
L	l		L		L	NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE
						QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMYTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHILLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVIFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEEPLTERIKYD PIPVRWNSLEGRLLRGYEQHANDGKDYISRN
						*DLRSWTAADMAAQITKRKWEAEEFAEQIKA YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	17.64	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	A	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQPGPPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPQQGDW EMYPVSWQTQESGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	A	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	1		in USSN	location	location corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq- uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Į	1	peptide	acqueiree	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
<u> </u>	-	<del> </del>		Soquence	<del> </del>	NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
	Ì					YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR
				1		RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
			1	ł	}	SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
		1				SLRTTGFYSGFSEVAEKRIKLLNNSDERLONS
						RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE
ļ						AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
						ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
[		[	ĺ	1	1	GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
ļ			}	1		VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE
			ĺ			LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
						KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
į		}	ł	ļ	}	QQKEGK
419	1769	Α	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
				į		NGVVLDTQQDQLENAKMEHTNASFDTFFCE
	J	ļ	ļ	1		TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP
		l .	1			PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
						VMGFHHVGQAGLELLTSGDLPALASQSARIT
			<u> </u>			GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS
			1			ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
1		i				LYLVSPLENEPKEMLTLSEYHERVRSQGQQL
						QQLQAELDKLHKEVSTVRAANSERVAKLVF
	ļ		ļ	1		QRLNEDFVRKPDYALSSVGASIDLQKTSHDY
			ľ			ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
			l	İ		GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP
l		,				PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ
		1		l		VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA
	ĺ	1	1	Ī	[	AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAOGPH
422	1772	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN
422	1//2	A	3412	2	421	YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
1	1	l	ĺ	İ	İ	PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
						QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
						VDTAGAPASPGPDVCE
423	1773	A	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
1	***	1 **	1 -120	1	1	
						RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG
•			i	1		RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
						I
						EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
						EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA
						EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QTTSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL  RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI
424	1774	A	3421	4	7688	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QTTSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA

SEQ ID	SEQ ID	Met	CEA	Predicted	Predicted end	Amino said converse (A - Aleria - C. C.
NO: of	NO: of	hod	SEQ ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1100		, , ,	location	
eotide	seq-	1	in	nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
1 1	uence		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq~ uence	uciice		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
			ŀ	residue of	Sequence	T=Threonine, V=Valine, W=Tryptophan,
				peptide	Sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i f			ĺ			/=possible nucleotide deletion, \=possible
			<b></b>	sequence		nucleotide insertion
l i						VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
1 1			ļ			GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
						VFVEENSPIGLAVARVTATDPDEGTNAQIMY
						QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
1						YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
1 1			}			LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
			1			DISDSLTYSFERGNELSLVLLNASTGELKLSR
						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
1						TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIQ
1						AVAATLATPPDHVVVFNVQRDTDAPGGHILN
						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
						LLTAISAQRVLPFDDNICLREPCENYMRCVSV
						LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
}						TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
} i			}	_		YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
1				_		TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
ļ l						QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
						VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
1 1						VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
						ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
}						LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
1 1						GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
						QASSLRLEPGRANDGDWHHAQLALGAIGGP
] !						GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
1 1						GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
l 1	i					SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
1 1						CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
						EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
1 1						RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
1						NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
1 1						SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
1 1						AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
j						LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
						NCTSITFSELKGFAERLQRNESGLDSGRSQQL
	.					ALLLRNATQHTAGYFGSDVKVAYQLATRLL
	.			j		AHESTQRGFGLSATQDVHFTENLLRVGSALL
] ]						DTANKRHWELIQQTEGGTAWLLQHYEAYAS
						ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
	}					FAGAKLPRYEALRGEOPPDLETTVILPESVFR
						ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
1 1						GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
	j					RPIINTPVVSISVHDDEELLPRALDKPVTVOFR
						LLETEERTKPICVFWNHSILVSGTGGWSARGC
						EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
[					l	NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
						RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
						DLPFACTVIAILLHFLYLCTFSWALLEALHLY
] ]						RALTEVRDVNTGPMRFYYMLGWGVPAFITG
		ļ		·		LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
<u> </u>					ŀ	VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1 1						PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
1 1	ł			l		FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
		- 1			i	LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
' -						PUCONDI OI DI UTI I IVOI TI IONI I INCLORI I ADC
						BI AUDIACIOS V CEI PELEBEVIS COPUS SUBBLE LE
					.	RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
					•	RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
ŀ				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
[	, ·	ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	-	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS
1	Ì		İ			TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
			1			EERLRENGDALSREGSLGPLPGSSAQPHKGIL
		İ				KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
}						GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
L						GTVDEDSSGSEFLFFNFLH
425	1775	A	3429	155	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS
						RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
1						GSAECLPGDPLGARRATRAHSPVPGPPPSLPA
1		ł	1			AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
1		į	1	ļ		CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
						SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV*
			'			KREFQRGPWAGMVILHRISAADPARAPGPDS
	1	1				NLQSALQQPATGCSEPAAVYSPPIGLWGA**P
		}			,	EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI
1	'				,	YELLENGQRAGTCVLEYATPLQTLFAMSQYS
			1			QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE
					, '	
1			]			EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS
426	1776	Ā	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
420	1770	A	3431	1002	309	SSOSLHCLSPROPNHYLOALRAVGGICODYD/
ĺ						SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
						AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
	1					KAGPHCSRLALTG\SHDFAINFDPENPECEGK
			1			RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST
İ				,		HPRPVFI\EISGVIASYRRCLPQIQLYGPTNVAP
İ						IINRVAEPAQREQSTGQATKYSVLLVLTDGV
İ						VSDMAETRTAIVRASRLPMSIIIVGVGNADFS
1						DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR
<b>;</b>						DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD
						VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI
	·					TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ
·			1			GISPGAPRPCTLATTPSPSP
427	1777	A.	3446 .	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
						GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
						ASRPEASGDCRAGRETAMATLEKLMKAFESL
				1		KSFQQQQQQQQQQQQQQQQQQQPPPP
1			]			PPPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPPPPPPP
						GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
						ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
						ESDVRMVADECLNKVIKALMDSNLPRLQLEL
				,	}	YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
						CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
						RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
						LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
			· .			KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
						TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
						LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
						AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
						ESRSDVSSSALTASVKDEISGELAASSGVSTPG
					<u>'</u>	SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
						TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
			'		-	GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
						FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
						VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
						APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
						SVKALALSCVGAAVALHPESFFSKLYKVPLD
L			ئــــــا	L		O TAN MARINE TO THE TANK THE ALPHA

No. of No. of No. of peptide ordine of the No. of U.S.N. o. o. o. o. o. o. o. o. o. o. o. o. o.	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
tence    Sequence   1944   1946   1944   1946   194		1	hod .	-		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence    09/496   correspond   ng to first		1					F=Phenylalanine, G=Glycine, H=Histidine,
unice	1	1 -					I=Isoleucine, K=Lysine, L=Leucine,
mmino acid reduced peptide sequence peptide sequence peptide sequence peptide sequence per sequ		uence				to last amino	
residue of peptide sequence peptide sequence seq	uence		'	914			
peptide sequence    Possible mueleotide insertion	Ī					• •	
INDICATE INTERPEDRY VSILLINYIDHIGD POVING ATALIC GTLICSIL SRSRPH VOD WMOTIRTLIGNTESI. ADCIPLLEKTIL MEDSSYNCKLACTA VENOVM SLCSSSYSELGI, QLIDVILTENSSYWLVRTEI. LETA ABIDFELVSTLEAK AKAPIHER AHHTYGI. LKLOERVILNYVIHLLGDED PRYRHVAAASI. RIVPKLLYKLOGGO ADPVAVAVARD SSYYL KILMHET OPPSHES VSTITRIY KGYPLLPSITD VTMENNILS RVIAAVS BELITSTTRALIT FOCCE. ALCLISTA FPVCIWSLOWHOCOYPPLS ASDESS. KSCTVORMATMIL TLIS SAWPPLDIS SANDESS. KSCTVORMATMIL TLIS SAWPPLDIS SANDESS. KSCTVORMATMIL TLIS SAWPPLDIS SANDESS. KSCTVORMATMIL TLIS SAWPPLDIS SANDESS. KSCTVORMATMIL TLIS SAWPPLDIS SANDESS. KSCTVORMATMIC TLIS SAWPPLDIS SANDESS. KSCTVORMATMIC TLIS SAWPPLDIS SANDESS. KSCTVORMATMIC TLIS SAWPPLDIS SANDESS. KSCTVORMATMIC TLIS SAWPPLD SANDESS. KSCTVORMATMIC TRIS SAWPPLD SANDESS. KSCTVORMATMIC TO SANDESS SANDESS SANDESS SAMPATK. QEEVWPALORAL PROPERTIES SANDES SANDESS SANDESS SANDESS SANDESS SANDES SANDESS SANDESS SANDESS SANDESS SANDESS SANDES						sequence	
TIEYPEEQYVSDILMYDHODPQVRGATAILC GTLICSLESSREPHVGSDWAGTRILTIGNTYSL ADCIPLLEKTIKDESSVTCKLACTAVENCYM SLCSSYSEJGJOLIDWITLRNSSYSYLYNTEL LETLAEDFELVSCLEAKAEM.HEGAHHYTGI. LKLGERVINNVYHLIGDEDPRYRHYAAAS. IRLVPKLPYKCDGQADPVVAVARDQSSVYL KLMHEDOPSHESVSTTIRTINGYGNLIPSITD VTMENNISRVAAVSHELITSTTEALTPGCCE ALGLESTAPPVCWSLGWHGGVPPLSASDESR KSCTVGMATMILTLESAWPFILDSAHDAAI LAGNILAASAPKSLESSWASEEANPAATK GEVWPALGBRALVSMVEQLFSHLAFSTOL AFVLDDVAGPARAALSPILTSPILSSPIRK GKEKEPUPQASVHLPSKYGSEASAASROSITS GPVTTKSSSSLSFYHLPSYLKJEDVLKATHA WYKVTDJQNSTEEFGGRASALDVLSQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGPGRASADDVISQILLE ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIASSGFFSREPMMATVC VQULKTLFGTNIASSGFFSREPMMATVC VQULKTLFGTNIASGFFSREPMMATVC VQULKTLFGTNIASGFFSREPMMATVC VQULKTLFGTNIASGFFSREPMMATVC VQULKTLFGTNIASGFFSREPMMATVC VQULKTLFGTNIAGROSTEVSVOJANTONICH STRIPTSTONICHOORYSTOLIC ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIAGROSTEVSVOJANTONICH STRIPTSTONICHOORYSTOLIC ATIQDIGKCVEELOYILKSCFSREPMMATVC VQULKTLFGTNIAGROSTEVSVOJANTONICH STRIPTSTONICHOORYSTOLIC NITSTYTRRADATATATATATATATATATATATATATATATATATA	1						
GTLICSILSRSRFIVGDWMGTIRTLTGNTFSL ADCPILLRKITLDGSSYTCKLACTAVRNCW SLCSSSYSELGLQLIDVILTLRNSSYWLVRTEL LETLALDFELVSTLEGGOADPYWAVARDSSYVL KLLOERVINNVUHLLGDEDPRYRHVAAASL IRLVPKLLPYCDGGADPYWAVARDSSSYVL KLLMEETOPPSHESVSTTIRTYRGYNLLPSITD VTMENNLSRVIAAVSSHEIDISTTRALTFGCCE ALCLISTAFPVCWSLGWHGGVPPLSASDESR KSCTVGGATMILTILSSAWFFLDLSSAHDDAL LLAGNILAASAPKSLRSSWASEEANPAATK QEEVWALADDALVSTENTERSHEINSTTRALTFGCCE AHVLDDVAFGAKAALPSILTPSTSPIRK GKEKERGEGASVFLSFKKGSBASASRQSDTS GPVTTSKKSSLGSFHLPSTYKLEDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLODIGKCVEELGYKSCSFSRFWMATVC VQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYTVCKSCSFSRFWMATVC VQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYTVCHASYTHETOPLALADA SLRNMVQAEQENDTSGWFDVLGKVSTQLKT NLTSVTKNRABKNAHMHILDFEPLYKLALQ YTTTTCVQLGKQVLDLLAQLVQLRVNYCLL DSQVYTGHAVTGHETOPLALADA SLRNMVQAEQENDTSGWFDVLGKVSTQLKT DSQVYTGHAVTGHETOPLALADA SLRNMVQAEQENDTSGWFDVLGKVSTQLKT NLTSVTKNRABKNAHMHILDFEPLYBALADA SLRNMVQAEQENDTSGWFDVLGKVSTQLKT VLTTTCVQLGKQVLDLLAQLVQLRVNYCLL DSQVYTGHAVAGTEFT GARGAGE LETTOKEVVVSMLIRLIQYHQVLEMFTLVLQQ GCHENEDENWKRLSRQJADILPMIA RQQMHI DSHEALGVLNTLFEILAFSSLRYDMLLRSMF VTPITMASVSTVQLWSGILLRIKVLISGSTED IVLSRIGELSFSPYLISCTVINRLRGGDSTSTLE HEISGKGKINLPEETFSRFLQLVGGILLEDIVT KQLKVEMSEQQHTFYCQGLLALGLDIVT KQLKVEMSEQQHTFYCQGLLALGLDIVT KQLKVEMSEQQHTFYCQGLLALLEDIVT KQLKVEMSEQQHTFYCQGLLALLEDIVT KQLKVEMSEQQHTFYCQGLLALLEDIVT KQLKVEMSEQGHTFYCQGLLALLEDIVT KQLKVEMSEQGHTFYCQGLLALLEDIVT KQLKVEMSEQGHTFYCQGLLALLEDIVT KQLKVEMSEQGHTFYCQGLGSSTSTLE HEISGKGKINLPEETFSRFLQLUGLLEDIVT KQLKVEMSEQGHTFYCQGLGSSTSTLE HEISGKGRAVTLTAVNTLLSPRYADDALSPS PYSSSHPLDGGGGSSTYLDSNLRA ARSMITTHPALLULFFRYLLARMODIL AGRIFERTAAATRIFRSRFLQLQVGLLEDIVT KQLKVEMSEQGHTFYCQGLAFAA ARGETLAANTGTVRACALLIFCDYVCQNLHDSE LTVLIVNHIDDLISL ENPPVQDFISCALLEDIVT KQLKVEMSEQGHTFYCQGLGSSTGLESDOLA ARLGMCREFYRRGALLIFCDYVCQNLHDSE HTTLLVHNIDGLISLENSFRVANDOL SPE PYSSSHPLDGGGGSTYLLEFRYLLTROODSTSTLE HISGKGRAVTLTAVNGTVQQLATCHARNSGVPA ANWSEFILLSAPCLEGGRAFATAAATRIFYCHPLLEAVAQQG LVSVSSTEFTVTHASLITYPVVQTLALAQU LVVVSSLFSH-HLIPPERKEDDIVKFVVATALLAQU VENDRASCHARMONAM WNSFFILSLAPCLSCHARMONA	<del></del>				Sequence		
ADCIPLIRATILADESSYTCKLACTAYRNOWN SLCSSSYSELGLIDIVALTRINSSYM_VETEL LETLAEIDERLVSTLEAKARNLIRGAHHYTICL LETLAEIDERLVSTLEAKARNLIRGAHHYTICL LETLAEIDERLVSTLEAKARNLIRGAHHYTICL LETLAEIDERLVSTLEAKARNLIRGAHHYTICL LETLAEIDERLVSTLEAKARNLIRGAHHYTICL KLIMETOPPSHERSYSTITRITRYGALPSTUL VTMENNLSRVIAANSHELITSTTRALTFGCCE ALCLLSTAFPVCUNSIGWHGGYPYLASADESR KSCTVGMATMILTLLSSAWFILDLSAHDOLA LAGNILAASAPKSLRSSWASEEBEANPAATK QEEVWPALGDRALVPMYGDLSSHLLKVINIC AHVLDDVAPPGARAALPELTPPSLSSPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASHAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGGASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSPKKOSEASAASROSISPIRK GKEKPFGCASVPLSTAMAKAROSINALPATASAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1	}					
SLCSSYSELGLQLIDVILTENSSYWLVERIE  LETLAEDPRIVESICARAEMHERGHHYTGL  LKLQERVLNNVVHILLGDEDPRVRIVAARDQSSVYL  KLLMEETOPPSIFESVSTITRYRGYNLLPSITD  VTMENNLSRVLAAVBELISTTRYTGYNLLPSITD  VTMENNLSRVLAAVBELISTTRYTGYNLLPSITD  VTMENNLSRVLAAVBELISTTRYTGYNLLPSITD  VTMENNLSRVLAAVBELISTRYTGGYEL  ALCLISTAFPVCIWSLGWHGVPPLSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  KSCTVGMATMIL TILSSAWPHLDSASDESR  GEVEVPALGDBALLAVPMVQLESBLLAVTNIC  AHVLDDVARGBAKAALPSILAVPSSLSPRKK  GKEKEPGPQASVPLSYRKOSEASAASRQSDTS  GPVTTSKSSSLGSFYBLFSYRLHDFULKATHA  NYKVTLDLQNSTEKRGGFRSALDVLSQILEL  ATLQDIKCVEEBLGVLKSCYRERPMATVC  VQQLLKTLGTTHLASQPDGLSSNYRSSQGAA  QRLGSSSVPRGLYNYCGMAYPHTTQALADA  SLRNMVQAGGENDTSGWFDVLQKVSTOLKT  NLTSVYTKNRAQANAHDHIRLPFTVLGALKQ  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL  DSDQVFIGFVLKQFYETVEQGFESSALPNIFF  FLVLLSYREYHSKQDIGFRUQLCDGIMASGR  KAVTHAPALQPVHDLFVLRGTNAAOKE  LETCKEVVSMLLRLQYHQVLEMFILVLQQ  GKEEPEKDKKKKLSRQADIDIPMLAKQQMHI  DSHBALGVUNTLFELLAFSSLRVVDMLLRSMF  VYPNTMASVSTVQLVISGILLAIRVISQSTED  IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE  BHSGGVGND-PETSFSRFLQLVGILLGOTHY  KQLKVEMSQQHTFYCQELGTLLMCLHIFKS  GMFRRITAATRILFRSDGCGGSFYLMGLISGTED  IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE  BHSGGVGND-PETSFSRFLQLVGILLGOTHY  KQLKVEMSQQHTFYCQELGTLLMCLHIFKS  GMFRRITAATRILFRSDGCGGSFYLMGLISGTED  IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE  BHSGGVGND-PETSFSRFLQLVGGLLEDIVT  KQLKVEMSQQHTFYCQELGTLLMCLHIFKS  GMFRRITAATRILFRSDGCGGSFYLMSON  AASGLFQAJQSRCENLSTTIMLKTLOCLGEBODLA  AKLGMCNEFVRRGALLLCOYVCQNLHUSE  LLTVINNHIDDLISSLSEPPYOPOBRAVERNS  AASGLFQAJQSRCENLSTTIMLKTLCDLCEG  HLSQSGAVLTLYVDFLLCTPREVDWYVHLVX  SCCWTRSDSALLEGAELVNERPADMAFM  MNSEPPLSLLAPCLSLGMSEGGCSALFFAA  AREVTLARVSGTVQQLTAVEHVPQPELPAP  AYWSLNDLFGDAALYQSLTALGHLRNNSGVA  PVSSHPLDQGHVSLETLARLAQVY  LVVSRLPSHLELPPSKERDVVYVATILEAL  SWHLIERGPLSLDLQAGDCCCLALQLFGL  WSVSSTEETVHACKLTLAQLERPLEVNYT TVPLVWLG  GESPPEEDTRATINVALVQAITSLVSLAGH  PVPSLSPATTAGLISHEKLLLQDIPPGWK	,						
LETLAEIDFRIVSTELAKAEMHIRGAHHYTCI LKLOGEVINNVUHLGDEDPRVRIVAAASL IRLYPKLFYKCDQGADPVVAVARDQSSVYL KLIMETOPPSHERVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELISTTRALTFGCCE ALCLLSTAFPVCINSIGWHGGYPPLASADESR KSCTVGMATMILTLLSSAWFILDLSAHGDAL LLAGNALAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVFWVSQLPSHLKVINIC AHVLDDVAPGPAIKAALPSLTYPFSLSPIRRK GKEKEPGGASVYFLSYTKKSGSAASRQSDTS GFVTTSKSSILGSTYHLFSYLKHDVLACHIAL NYKVYLDLQNSTEKFGGGRASALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQULKTLFGTHLAGSPGGLSSSTSGQFA QRLGSSSVRPGLYTYCTMAPYTHFTQALADA SLRNWQAGEDBINTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHHHIRLFEFLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQTAVYCLL DSDQVFIGFVLKQFEYEVGGFRESAHPNIFF FLVLLSYSEYHSKOGIGFKIGLOGHASGR KAVTHAPALQPIVHDLFYLKGTNKADAKGE ETCKEVVVSMLLRLIQYHOVLEMFLVLQQ CHKENEDKWKLSRQLADILPMLAKQOMHI DSHEALGVANTHELAPSLRFVNLRSMFLVSHW VPPNTMASVSTVQLWSGLALIR VVLISQSTED VLSSIGLSFSYLVLSCTVNRLRDGDSTSTILE EHSEGKQKNLPEETFSRFLQCVGILLEDIVT KQLKVVEMSEQQHTFYCQCLGTLLMCLHIFKS GMFRITAATTALFRSDGGGSFTLDSINLIK ARSMITHFALVLLWCQLLLVNHTDYNWW ARVQGTKRHISSTSTLLSPQNNSCHSHVRIN AASGLFIQAIQSRCENLSTTMLKKTLQCLGGI HLSQSGAVLTLYVDRLLCTFRVLLDSINLIK ARSMITHFALVLLWCQLLLVNHTDYNWW ARVQGTKRHISSTSTLLSPQNNSCHSHVRIN AASGLFIQAIQSRCENLSTTMLKKTLQCLGGI HLSQSGAVLTLYVDRLLCTFRVLARMYDIL ACRRVEMLAANLQSSMQLPMEELNRIQGS LQSGAVLTLYVDRLLCTFRVLARMYDIL ACRRVEMLAANLQSSMQLPMEELNRIQGS LQSGAVLTLYVDRLLCTFRVLARMYDIL ACRRVEMLAANLQSSMQLPMEELNRIQGS LSGSAVLTLYVDRLCTFRVTAATHOLIK SWYNSSTEPTIBLAGALLARNINGNY SQLVFRSBSLLBAGALLARNINGNY SQLVFRSBSLLBAGALLARNINGNY SQLVFRSBSLLBAGALLARNINGNY ANSENLSLLAPCLSLGMSEISGGGKSALFS PPVSSIPLLBGGGHALLYNNTFRVPSLAPRING ASSGLFQALGSCCCLALQPGLWYREITHAA SWYSLADDIRGDAALYQSLFTLARALQQY LVVSKLPSHLHEIPPEKEDIVKYPTLARMYDIL ACRRVEMLAANLQSSMQLPMEELNRIQGS GQLLSFERRTNITRALFLYNNSTRYPPLYWKLG WSPKFGGBFGTAFEPEIPPETIQRKEVPEEPTRP AYWSLADDIRGDAALYQGLPAHPLYNNTRYPPLYWKLG WSPKFGGBFGTAFEPEPPETIQRKEVPEEPTRP AYWSLADDIRGDAALYQGLPAHPLYNNTRYPPLYWKLG WSPKFGGBFGTAFEPEPPETIQRKEVPEEPTRY INTLGWYRCGGEGGGALTRICTHARNINGYPA ACMOTASNICLEGGGGCGALFS							
IRLUPKLIFYKCDQGQADPVVAVARDQSSVIL KLIMETOPPSHERVSTITRINRGYNLLPSHID VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCINSIGWHGGYPLASNDESR KSCTVGMATMILTLLSSAWPILDLSAHQDAL LLAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVZQLPSHLKVINIC AHVLDDVAPGPAIKAALPSLTVPPSLSPIRRK GEKEPGPGASVPJKKSGSAASARGSDTS GPVTTSKSSSLGSTYHLPSYLKHDVLASTHAL NYKVYLDLQNSTERFGGGRASALDVLSQILEL ATLQDIGKCVEELIGYLKSCPSKEPMMATVC VQULKTLFGTHLASQPGGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNINVQAEQENDTSGWPDVLQKVSTQLKT NLTSVTKNRADKNAHHHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQTAVYCLL DSDQVFIGFVLKQFEYTEVGGFRESEAHPNIFF FLVLLSYREYHSKOJGPKSLGGNASGR KAVTHAPALQPIVHDLFVLGTNKADAGKE ETGQKEVVSMLLRLIQYHOVLEMFULVQQ CHKENEDKWKLSRQLADHLPMLACQOMHH DSHEALGVANTLFELLAPSLRPVDMLRSMPULVQQ CHKENEDKWKLSRQLADHLPMLACQOMHH DSHEALGVANTLFELLAPSLRPVDMLRSMP VTPNTMASVSTVQLWSGLALRVLLSQSTED VLSRGLSFSFYLLSCTVNRLBODDSTSTLE EHSEGKQKNLPEETFSRFLQLVGILLEDIVT KQLKVVEMBEQQHTPYCQELGTLLMCLHHIPKS GMFRITAATATLRFSDGGGSFYLDSLNLR ARSMITHFALVLLWCQBLLVNNTDYRWW ARVQQTKRHISLSTKLLSPQMOGEEDSDLA ARLGMCNRETVRRGALILCOYVCQNLHDSE HLTWLNNHIQDLISS.SLEPPYQDGTSTLDSLNLR ARSMITHAFALVLLWCQBLLVNNTDYRWW ARSQFTKAHGSSTKLLSPQMSGEEDSDLA ARLGMCNRETVRRGALILCOYVCQNLHDSE HLTWLNNHIQDLISS.SLEPPYQDGTAVHRNIN AASGLFQAJGSRCENLSTTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFRYLARMYDLL ACRRVEMLAANLQSSMQLPMEELNRIGES CAGGRAVLTLYVDRLLCTFRYLARMYDLL ACRRVEMLAANLQSSMQLPMEELNRIGES ASGLFQAVLTLYVDRLLCTFRYLARMYDLL ACRRVEMLAANLQSSMQLPMEELNRIGES AYSHLDPGDAALYQDLATARAMYDL ACRRVEMLAANLQSSMQLPMEELNRIGES SWYSSIPLJBGGGHVSLTVSPDKDWYVHLVX SQCWTRSDSALLBGALLARNDSTYLE SWHLHEQPIPSLDLQGABCCCLAIQLPGL WSVYSSTEPVTHACSLICAGENGTRAPPIPVCWPLARMYDL ACRRVEMLAANLQSSMQLPMEELNRIGES FYSSIPLJBGGGHVSLTVSPDKDWYVHLVX SQCWTRSDSALLBGALLVNNTRYPOPPLAME GESPPEEDTERTQINVLAVQATISLUTRPGK SWYSSTEPVTHACSLICAGENGTRAPPIPVENCHUR GESPPEEDTERTQINVLAVQATISLUTRPGK LSWGLSPFEEDTERTWALLGGVYCQTAVHTERAVALICAGENSG							
KLLMHETOPPSHFSVSTITRIVRGYNLIPSITD VTMENNLSRVLAAVSHELITSIVRGYNLIPSITD VTMENNLSRVLAAVSHELITSIVRALTIFGCCE ALCLISTAFPVCIWSLGWEGGYPLASABESR KSCTVGMATMILTLISSAWPLDLISAHGDAL ILAGALLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPWEQLFSHLKVINIC AHVLDDVARGEAKAALPSINPPSLSPIRRK GKEKERGQASVPLSPKKGSEASAASGDTS GPVTTSKSSLGSFYHLPSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELGYLKSCFSREPMATVC VQQLKTLFGTNLASQPDGLSSNPSSSQGRA QRIGSSSVRGCHLGYKYCFMPYTHFTOALADA SLRNMYOADDENDTSGWPDVLOKVSTOLKT NLTSYTKNRADKAIENHRIFPELVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVEIGFVLKQFSEVELOGPESEALPPSITE FLVLLSVERYHSKQIIGPKIUQLCDGIMASGR KAVTHAPALQPVHDLFVLROTIKAADAGKE LETQKEVVSMLIRLIQVHQVLEMFILVLQ CHENEDEWKRLSRQJADILIPMLAKQOMHI DSHEALGVLNT-HEELAPSSLRAPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRQGLSSSPYLISCTVINRIRDGDSTSTLE EHSEGKQIKNLPETFSRFILQLVGILLEDIVT KQLKVEMSEQOHTFYCQELGTLLMCLHIFFSK GMFRRITAAATRLFSBGCGGSFYTLDSILNR ARSMITHPALVLWCQILLLVHITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNETFURGALLVLWCGILLLVHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNETFURGALLFULWCOLLLDIVT KQLKVEMSEQOHTFYCQELGTLLMCLHIFFSK GMFRRITAAATRLFRSBGCGGSFYTLDSILNR ARSMITHPALVLWCQILLLVHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNETFURGALLFULWCOLLLDIVT KQLKVEMSEQOHTFYCQELGCUCOLLDOIL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LOSSGLAQRIRGRYSTLSTTMLKKTTLCLEGI HLSQSGAVILTYVDRLICTFRVLARMVDL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIRGRYSTLSLEGHAMVDL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIRGRYSTLSLEGHAMVDL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIRGRYSTLSLEGHAMVDL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIRGRYSTLSLEGHAMVDL ACRRYEMLAANLQSCHAPEDMANPM MNSEPNLSLLAPCLSLGMSEEGGKSALFAA AREXTILARVSGLAGGALLGGALLVARFADMANPM MNSEPNLSLLAPCLSLGMSEEGGVALGHKRNSGVPA FLTTLLRNIBLARLEVNSYTYPHLVWKLG WSWRGGGFGTAFFETWETLGEKSVFKETYX INTIGWTSRTGFETWATLGCHESTVETHTY INTIGWTSRTGFETWATLGCHESTVFKETYX INTIGWTSRTGFETWATLGCHESTVFKETYX INTIGWTSRTGFETWATLGCHESTVFKETYX INTIGWTSRTGFETWATLGCHESTVFKETYX INTIGWTSRTGFETWATLGCHAMTHAITYGAWD PYPSLSPATTGALISHEKLLLUPREELGSMS	1						LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
VTMENNLSRVIAAVSHELISTTRALITGCCE ACCLLSTAFPCVWSLGWIGGVPLOSASDESR KSCTVOMATMILTLLSSAWFILDLSAHQDAL ILAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVEQLFSHLIKVINIC AHVLDDVAFGRAKAALPSLTNPPSLSPIRR GKEKFIGGQASVYLSPKKOSEASAASRQSDTS GPVTTSKSSSLGSYYHLPSYLKHDVLKATHA NYXVTLDLONSTEKFGGFLRSALDVLSQILEL ATLODIGKCVEELLGYLKSCFSREPMMATVC VQQLKTLFGTINASOFDGLSSAASARGSDTS GPVTTSKSSSLGSYYHLPSYLKHDVLKATHA NYXVTDLONSTEKFGGFLRSALDVLSQILEL ATLODIGKCVEELLGYLKSCFSREPMMATVC VQQLKTLFGTINASOFDGLSFSALDVLSQILEL ATLODIGKCVEELLGYLKSCFSREPMMATVC VQQLKTLFGTINASOFDGLSFSSAGGRA QRLGSSSVRGUHYYCFMAPYTHFTQALADA SLRINMYQABCBDTISGWFDVLQKVSTQLKT NLTSYTKNRADKNAEDHIRLFEPLVIKALKQ YTTTTCVQLKQVQLLAQLVQLXVNYCLL DSDQVFIGFVLKQFSYEEVGPRESEAILPNIFF FF.VLLSVERYHSKQNIGHKIQLCDGIMASGR KAVTHAPALQYHDLIPLYLGTNKADAGKE LETQKEVVVSNLLRLIQYHQVLEMFILVLQQ CIKENDEWKRKLSRQIADILINAGNGMH DSHEALGYLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGIALINLGSQTED IVLSRQIGLISFSYVLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPETFSRELQVGLILEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRATTAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLWCQULLLVHNTTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGCEEDSDLA AKLGMCNRETVRGALLFCPVQCHOLIBDIV KQLKVEMSEQQHTPYQCPLHDSE HLTWLIVNHQDLISLSHEPVYQDFISAVHRNS AASGLFQAIQSGCENLSTPTMKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVEMLLAANLQSSMAQLYQCHNDSE HLTWLIVNHQDLISLSHEPVYQPFISAVHRNS AASGLFQAIQSGCENLSTPTMACTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVEMLAANLQSSMAQLYPELINRIQEY LQSSGLAQRHQRLYSLTLDFFLSTMQDSLSPS PPVSSHDLDGDGMYSLETVSBELDVYPHUNKLY SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLAPCLSLOMSESGGGKSALFFA AREVTLARVSGTYQLAVHYPQFELPAEP AAYWSKLNDLFGDAALYQSLPFLARALQY LVVVSKLPSHLHPPEKKEDVKFVVATLEAL SWHLHBQIPLSLDLQAGLDCCCLALQLPGL WSVVSTEFVTHACSLIVCVHFLEAVAVQPG EQLLSPERTNTPKASSEEGECVPPTTQPKYT TAACENVAEMKYBSLOSVLALGHKRNSGVPA FLTPLLENNISLARLPLYKATTRYBLSAMT VPVAGNPAVSCLEQQFRNKPYKALDTFGKK LSIRGIVEGEIQAMVSKRENIATHHLYQGWWD PVPSLSPATTGGLSHEFLIGSMYS LSIRGIVEGEIQAMVSKRENIATHHLYQGWWD PVPSLSPATTGGLISHEKLLLQHAWDFLGGMWD	1						IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
ALCLISTAFPYCTWSLGWHCGVPPLSASDESR KSCTVGMATMILTLISSAWPLSAHDOAL ILAGALLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMEQETALLAVINIC AHVLDDVAPGPALKAALPSLTNPPSLSPIRK GKEKPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQUEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRFQLYNYCFMAPTHTTQALADA SLRNMYQAEQENDTSGWFDVLQKVSTOLKT NLTSYVKNRADKNAIFMHRILFPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVYYCLL DSDQVFIGFVLKQFEYEWCQFRESEAIPNIFF FLVLLSVERYHSKQNIGPKIQLCDGIMASGR KAVTHAPALQPVEIDLFVLRGTNKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVJQ CHKENBEDK WKRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRFYDMLLRSMF VTPNTMASVSTVQL WISGILALRVLISQSTED IVLSRQLSISPSPYLISCTVINRLDGDSTSTLE EHSEGKQKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQGHTYCQCELGTLLMCLHIFKS GMFRRTAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVHITDYRWW AEVQOTPKRHSLSSTKLLSPQMSGGEEDSDLA AKLGMCNREIVRGALLLVHWQGSGEEDSDLA AKLGMCNREIVRGALLSTLYLDWGCGEEDSDLA AKLGMCNREIVRGALLSTLYLDWGCHLSUSLNLR ARSMITHPALVLLWCQILLLVHNTDYRWW AEVQOTPKRHSLSSTKLLSPQMSGGEEDSDLA AKLGMCNREIVRGALLFULWGCHLGHFKS GMFRRTAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVHRTILYRNS AASGLFJQALGSGENSTSTLDSTLYLDWGLHGEN LAGGCNREIVRGALLLFULWGTLLTPRVLARMVDL ACRAVEMLAANLQSSMAQLPMELNRIQEY LQSSGLAQRRIQRLYSLDERFRLSTMQDSLSP PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRISSALLBGABLVNRPAEDMNAFM MNSEPNLSLLAPCLSLGMSEGSGGKSLFFA AREXVILARVSGTVQLPAVHLYRGPLLAPGL LSWHLHEGDFLSLDQAGLDCCCLALQLPGL WSVVSSTEFVHACSLLYCVHILLBAVAVQPG EQLLSPERRTNTPKAISEEGELVPNTOPKYI TAACEMYAEMVSSLOPENTSPLYNKLI WSPKRGGDFGTAFPETVEFLQEKSVKRETYR INTIGWTSRTGGETWAILLGVNLTPGFWK LSWHLHEGDFLSLDQAGLDCCCLALQLPGL WSVVSSTEFVHACSLLYCVHELBAVAVQPG EQLLSPERRTNTPKAISEEGEEVPNTOPKYI TAACEMYAEMVSSLOPKARVSLAATHHILYQAWD PVPSLSPATTGGLETWALDGLWATHFGKK LSIRGIVEQEIQAMVSKREVIATTHHILYQAWD PVPSLSPATTGGLAUSHELGSLMATHHILYQAWD PVPSLSPATTGGLISHEKLLLUGNFERGGMSM							
KSCTVGMATMILTILISSAWFFLDLSAHQDAL ILAGNILAASAFSLIRSSWASCEANPAATK QEEVWPALGDRALVPWEQLFSHLKVINIC AHVLDDVAPGAIKAAJPSLTNPPSLSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSYHLPSYLKHDVLKATHA NYKVTDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELGYLKSCFSREPMMATVC VQLLKTLFGTINLASOFDGLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQALADA SLRNMYQAEQENDTSGWFDVLOKVSTQLKT NLTSVTKNRADKNAIRNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLQVRNYCLL DSDQVFIGFVLKOFEVEVGGPESEAUPNIFF FLVLLSYERYHSKQIIGPKIQLCDGMASGR KAVTAPALQPTVBDLFVLKGTNKADAGSGR KAVTAPALQPTVBDLFVLKGTNKADAGSGR KAVTAPALQPTVBDLFVLKGTNKADAGSGR KAVTAPALQPTVBDLFVLKGTNKADAGSGR KAVTAPALQPTVBDLFVLKGTNKADAGSGR VPTPTTMASVSTVQLWISGIGLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGIGLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGIGLAVLGSQTSD IVLSRIQGLSFSPYLISCTVINKIRDGDSTSTLE EHSGKQKNLPEFTFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMRRITAAATRLFRSDGGGGTLUVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMRRITAAATRLFRSDGGGGTTLLMCLHIFKS GMRRITAAATRLFRSDGGGGTTLUNGLHIFKS GMRRITAAATRLFRSDGGGGTTLUNGLHIFKS GMRRITAAATRLFRSDGGGGRSTLDSLUNG ARSGIFTQAIGSGCENLSTPTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLAANLQSSNAQLFILKSLTQLCEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLANLANLOSSNAQLFILKSLTQLCEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLAANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLAANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLANLANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLAANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLANLANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLANLANLOSSNAQLFILKSLTQLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACREVEMLANLAGLSLAGLANGLANGLANGLANGLANGLANGLANGLANGLANGL	1						
ILAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMQLTSHLLKVINIC AHVLIDDVAPGPAKAALPSI.TNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSTYHLEPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQULLKTLFGTNLASQPTGLSSNSSKGGAA QRIGSSVVPLSPGLYHYCMAPVTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYKINADAKNAHNHEPEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNTCLL DSDQVFIGFYLKQFEYTEVGGFRESALIPNIFF FIVLLSYRSYHSKQIGTBKIQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LEPTQKEVVYSMLIRLIQYHQVLEMFILVLQQ CHEENEDKWRLSRQIADDILPMLAKQOMHI DSHEALGVINTLFELLASPRYDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRQGELSFSPYLISCTVINRLRGGDSTSTLE EHSGKQKINLPEETTSRFPLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCHHFKS GMFRRITAAATTLFRSDGCGSSYTLDSINLR ARSMITHPALVLLWCQULLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSGMSGEESDLA AKLGMCNRETVRRGALILFCDYVQOLHDSE HLTWLIVNHIODLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTI,VVRRLCTFFRVLARMVDIL ACRRVEMILAANLQSSMAQLFMEEINGPS LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLOGGHYSLTVPDKLWTVDKLY SQCWTRSDSALLEGAELVNRPAEDMAPAM MNSSFNLSLLAPCLSLGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDGHYSLETVSPKDWTVHJVK SQCWTRSDSALLEGAELVNRPAEDMAPAM MNSSFNLSLAPCLSLGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDGHYSLETVSPKDWTVHJVK SQCWTRSDSALLEGAELVNRPAEDMAPAM MNSSFPLSLLAPCLSLGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDGHYSLETVSPKDWTVHJVK SQCWTRSDSALLEGAELVNRPAEDMAPAM MNSSFPLSLAPCLSLGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLDPLGFGGAALLGHKRNSGVPA FLTPLLRNIIISLARLPLNNSTRVPPLVWLIG WSVYSSTEFTVHACSLTLYCHFLEAVAVQPG EQLLSPERTNTFRAISEEEEEDPTNPTNPKYI TAACEMVAEMUSSLOQGHACCCLALQLPGL WSVYSSTEFT THACKTLLQLPGLUFFLUWEL SVERPEDDFGTAFPEIPTSTLQEKEFTIR NITLGWTSRTQFEETWATLLLGVLVTQPLVME QESSPPEDTERTGNINVLAVQATISSLJSAMT VPVAGRPAVSLCQQPRKRLKALDTRFGRK LSIRGIFVSGEIQAMVSKRENATHHI-QAWD PVPSLSPATTGALLSHELLLQNPETELGSMS	1						
QEEVWPALGDRALVPMCQLESHLLKVINIC AHVLDDWAPGARKAJPSI.TSPPSLSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSTYHLFSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFRAALDVLSQUEL ATLQDIGKCVEELI.GYLKSCFSREPMATVC VQQLLKTLFGTNLASQPFGLSSNFSKSQGRA QRLGSSSVRPGI.YHYCMAPYTHETQALADA SLRNMYQAGENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYTEVGGPRESEAIPNIFF FVLLSYERYHSKQIGFPKIQLCOGIMASGR KAVTHAPALOPIVHDLFVLKGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEAHIVLQQ CHKENEDKWKRLSRQLDIPMLAKQQMHI DSHALGVLNTLFEILAPSLRPVDMLLRSMF VTRNTMASVSTVQLWISGILALBVTLSQSTED IVLSRQELSFSPVLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETTSRFLLQLVGGILLEDIVT KQLKVEMSEGVAFYVSCOLGTLLMCHHFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQJLLLVHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCKRETVRRGALLHCDYVCQHLHDSE HLTWLINVHIDQLISLSTEPPVQDFTSAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLTQCLEGI HLSQSGAVLTLYVDRLLTQCLEGI HLSQSGAVLTLYVDRLLTCRVLAKWODL ACREVEMLLAANLQSSMAQLPMEELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPP PPVSSHPLDGDGHYSLETVSPKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMAPM MNSEPNLSLLAPCLSLGMESISGGQKSALFEA AREVTLARVSGTVQQLPAVHHYFOPELPAEP AAYWSKLNDLFGDAALYSCPPLAPRALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEPVTHAGSLIFVCHFLEKAVQPG EQLLSPERTNTFKAISEEEEEEVDFNOPKYV TAACEMVYBAUFSLORGVLAGHKRNSGYPA FLTPLLRNIIISLARLPLNSSTRVPLWKLG WSVPSGGGGTAFPEIPVETLQREKEFIYR INTLGWTSRTOPEETWATLLQNLYTOPLWME QEESPPEEDTERTQNIVLAVQAITSLLSAMT VPVAGRPAVSCLEQQPRKKLLADTFFCRK LSIRGIFVEGIGQMVSKRENNATHHILQAWD PVPSLSPATTGALLSTVALLDGPRELGSMS	1						l · · · · · · · · · · · · · · · ·
AHVLDDVAPGPAKAALPSI.TNPPSLSPIRKK GKEKEPGGASVPLSPKGASAASQSDTS GPVTTSKSSSLGSPYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQULKTLFGTHLASQFPGLSSNFSKGGRA QRIGSSSVPGGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYKINADKNAHNHISPEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYYEVQGFRESEAILPNFF FLVLLSYBRYHSKQIGFBYIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVYSMLTLIQVHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQOMHI DSHEALGVINTLFEILASPRYDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLLQSTED IVLSSIGGLSFSPYLISGTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGSFYTLDSLNLR ARSMITHPALVLLWCQLLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSFQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQHLDSE HLTWLIVNHIODLISLSIEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFRVLARMVDIL ACRRVEMLLAANLQSSMAQLFMEELNRIQEY LQSSGLAQRHQRLVSLLDRFRLSTMQDSLSPS PPVSSHPLJOGGHSVLTVPDKLWRDDSLSPS PPVSSHPLJOGGHSVLTVPDKLWRDDSLSPS PPVSSHPLJOGGHSVLTVPDKLWTDGLSSFGGKSALFEA AREVILARVSGTVQQLPAVHHYFOPELPAEP AAYWSKLNDLGGHSLJAPCLSCGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFOPELPAEP AAYWSKLNDLFGDALJVQSLPTLARALAQY LVVVSKLPSHLHLPPEKERDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFFVHACSLITVCHFILEAVAVQPG EQLLSPERTINTFKAISEEEEEVDPTNQPKYI TAACEMVAEMVSLGQVTRYLGKKRNSGVPA FLTFLLRNIISLSARLLPLNNSYTRVPDLOWFIN TAACEMVAEMVSLCQSVTRLAGKKRNSGVPA FLTFLLRNIISLSARLPLNNSYTRVPDLOWFIN TAACEMVAEMVSLCQGFRKVLKALDTFFCFK LSIRGIVEQEIQAMVSKRENNATHH VQAWD PVPSLSPATTGALLLGUFFRELGKES	1						l .
GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSLGSFYHSYLLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCTSREPMMATVC VQLLKTLFOTNLASQPGGLSSNPSKSGGRA QRLGSSSVRFGLYHYCFMAFYTLFTQALADA SLRNIWQAGEENDTSGWFDVLKSTOLKT NLTSVTKNRADKNAIFHHIRLFEPL-VIKALKQ YTTTTCVOLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFYEVGGFRESEAIIPNIFF FLV1LSYERYHSKQHIGIPKHIQLCDGIMASGR KAVTHAPALQPIVHELYRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFLVLQQ CHKENEDKWKRLSRQIDLIPMLAKQOMHI DSHEALGVLNTLFEILAFSLRFVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQURNLPEFFSRFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQCHGTLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYLDSLNLR ARSMITTHPALVLLWCGLGLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTIPMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARWVDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAGRIGNLYSLLCTFRFVLARWVDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAGRIGNLYSLLCTFRFVLARWVDIL ACRRVEMLSLAPCLSGLESVENDYVHLVK SQCWTRSDSALLERLSTNGDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGLELVNRIPAEDMNAPM MNSEFNLSLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVQPELPAEP AAYWSKINDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLH.PPEKEKDIVKFVVATLEAL SWHLHEDPIPLSDLQAGLDCCCLALQLPGL WSVVSSTEFVHACSLIVCVHFILEAVAVQPG EQLLSPERRINTPKAISEEEEEVDPNTQWPKYI TAAACEMWAEMVESLQOPRINFLKALDTRFGRK LSURGIVERGEMS SYFKPGGFGTAFPEPTLQEKEVKKERFIYR NTILGWTSRTQFFETWATLLGVLYTOPLVME QEESPPEEDTERTQNVLAVQAITSLVLSAMT VPVAGNPAVSCLEQOPRINFLKALDTRFGRK LSURGIVEQEQAMVSKEENATHHLYQAWD PVYSLSPATTGALLSTEKSMS	1						3
GPVTTSKSSILGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGAA QRLGSSSYRGLYHYCPMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGGFRSESAIINNIFF FLVLLSYBEYHSKQIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVEDLFVLRGTNKADAGKE LETYGKEVVSMLLRLIQHYQLEMFILVLQQ CHKENEDKWKRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRIQELSFSYLISCTVINRLRDGDSTSTLE EHSEGKQUKNLPEETFSFILLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLDLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLDVONGEEEDSDLA AKLGMCNEIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIOSNCENLSTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRIYSTPIMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRIYSTPIMLKSTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRIYSTPIMLKSTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRIYSTDRFRISTMQDSLSFS PPVSSIPLDGDGHVSLETVSPROWVYHLVK SQCVTRSDSALLEGGALVNRIPAEDMNAFM MNSEPNLSLLAPCLSIGMSEISGQKSALFEA AREVITLARVSGTVQQLPAVHHYOPELPAPP AAYWSKINDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPFEKEKDIVKFVVATLEAL SWHLIEBGPUSLDLQAGLQCCLALQLPGL WSVVSTEFVTHACSLLVCVPHILEAVAVQPG EQLLSPERTNTPKAISEEEEVDPNTQNPKYI TAAACEMVAEMVESLGSVLALGHKRNSGVPA FLTPLLRNIISLAARLPLVNSYTRYPLVWKLG WSPKPGGFGTAPPETPEFLQKEVFKEFTYR NTLGWTSRTOFEETWATLLGVLVTOPLVME QEESPPEEDTERTQNTVKALGALTRFGRK LSIRGIVSQEQAMVSKEENIATHHLYQAWD PVYSLSPATTGALLSIGMS	1						—·· ;
NYKVTLDLQNSTEKFGGFIRSALDVLSQUEL ATLODISC VOED 164 KCYESTERPMAATVC VQQLKTLFGTNLASQFDGLSSNPSKSQRA QRLGSSSVRFGLYHYCFMAFYTHFTQALADA SLRNMVQAGENDTSGWFDVLQKVSTQUKT NLTSVTKNRADKNAIPHHIRLFEPLVIKALKQ YTTTTCVOLQKQVLDLQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAIIPNIFF FLV1LSYERYHSKQIIGFKIIQLCDGIMASGR KAVTHAIPALQPYHEULYRGTNKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAFSLRFVDMLLRSMF VTPNTMASVSTVQLWISGILALIRVLISQSTEID IVLSRQGLSFSFYLISCTVINRLRDGDSTSTLE EHSBGKQIKNLPEFTFSRFILQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIIFKS GMFRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVHITDYRWW AEVQQTFKRHSLSSTKLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAJQSRCENLSTITMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLANILQSSMAQLPMEELNRIQEY LQSSGLAQRIGRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLBGASLVNRIPAEDMNAFM MNSEPNLSLLAPCLS GMSEISGGGSSAFFA AREVILARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALLYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILAVALQPF LTPLLRNIISLARLPLVNSYTRVPLLWKLG WSPKFGGGFGTAFFEPVEFIQEKEVKKEFTYR NTILGWTSRTQFEETWATLLGLYVPLLKEFTYR NTILGWTSRTQFEETWATLLGLYVPLLVMCLGF SYSPKGGDGFGTAFFEPVEFIQEKEVKKEFTYR NTILGWTSRTQFEETWATLLGLYVPLVMLGAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEQAMVSKRENLATHLYQAWD VPVSKSPATTGALLSTREKLLLQNSYRCHLANDT	1.						
ATLODIGKCVEELIGYLKSCFSREPMMATVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCPMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHHIRLFEFLVKALKQ YTTTTCVQLQKQVLDLLAQLVQURVNYCLL DSQVFIGFVLKQFFEVEVGOFRESEALIPNIFF FLVLLSYERYHSKQTIGIPKIQLCDGIMASGR KAVTHAIPALQPTVHDLPVLRGTTNKADAGKE LETQKEVVSMLLRLQYHQVLEMFLVLQQ CHKENEDKWKRLSRQIADILPMIAKQQMHI DSHBALGVINTLFEILAFSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED INLSRIQELSFSPYLISCTVINRLBSGDSTSTILE EHSEGKQUNLPEETFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYLQUSGGSFYTLDSLNLR ARSMITTHPALVLWCQULLVNHTDYRWW AEVQQTKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALLFCDVVCQNLHDSE HLTWLIVNHODLISLSHEPPVQDFSAVHRNS AASGLFQAQISGCENLSTFTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRRQRLYSLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MISSENLSLAPCLS GMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHNVFQELPAEP AAYWSKLNDLFGDAALVGSLFTLARALAQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHERPIPLSDLQAGLDCCCLALQLPGL WSVLYSTEFVTHACSLIVCVHFILEAVAVQPG EQLLSPERTNTPK ALSEEEEVDPNTQNPKYI TAACEMVAEMVESLGSVLALGHKRNSGVPA FLTPLLRNIISLAARLPLVNSYTRVPLVWKLG WSPKPGGDFGTAPFPETVEFLGKEVFKEFTYR NTLGWTSRTOFFETWATLLGYLVTQPLVME QEESPPEEDTERTQNTVAQAGITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEQAMVSKRENIATHLYQAWD PVPSLSPATTGALLSTEKLLLQNFRELEGSMS	1.	]			·		
ORLGSSSVPGLYHVCFMAPYTHFTQALADA SLRNMVQAGENDTSGWPDU,QKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVPIGFVLKQFEVGOFRESEAIIPNIFF FLVLLSYERYHSKQDIGPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWRLSRQIADIILPMLAKQQMFI DSHEALGVLNTLFAPSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCOELGTILMCLHIFKS GMFRRITAAATRLFRSDGCGSFYTLDSLNLR ARSMITTHPALVLLW CQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSFQMSGEEDSDLA AKLGMCNREFVRGALILFCDYVCOPILHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQFRQRLYSLLDTFRISTMQDSLSFS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNISLLAPCLS GMSEISGGKSALFFEA AREVTLARVSGTVQQLPAVHHVPQPELPAEP AAYWSKINDLFGDAALYQSLYTLARALAQY LVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLHBQDFISLDLQAGLDCCCLALQLPGL WSVVSSTSFTHACSLIYCVATLEA SWHLHBQDFISLDLQAGLDCCCLALQLPGL WSVVSSTSFTHACSLIYCVATLEA SWHLHBQDFISLDLQAGLDCCCLALQLPGL WSVKSTSHTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNITKAISEEEEEVDPNTONPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFIEVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGLYNTYPPLWKLG WSPKPGGDFGTFIEVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGLYNTYPPLWKLG USSPREDTERTGNAVAVAQATTSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALLSHCLLQNPTERGLGSMS	1						· <del>_</del>
SLRIMYQAEQENDTISGWPDULQKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVJDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAIIPNIFF FLVLISVERYHSKQIPKIQLCOGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRYDMLLRSMF VTPNTMASVSTVQLSGIALIRVLISQSTED IVLSRQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQIKNLPEETTSRFLLQVGILLEDIVT KQLKVEMSEQQHTPYCQELGTLLMCLHIFFKS GMFRITAAATRLFRSDGCGGFYTLDSLNLR ARSMITTIPALVLLWCGILLIVHTYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDISDLA AKLGMCNREIVRRGALLIFCDYVCQNLHIDSE HLTWLINVNHIQDLISHFPYQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSFS PPVSSHPLDGGGHVSLLDVFRLSTMQDSLSFS PPVSSHPLDGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNIPAEDMNAFM MINSENISLLAPCLSLGMSEISGG(KSALFEA AREVILANVSTCVQLYAVHHVPGELPAEP AAYWSKLNDLFGDAALVQSLPTLARALAQY LVVVSKLPSH.HILPFEKKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVHACSLICVFFLEXAVATVEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVHACSLICVFFLEXAVAYPA FLIPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFEIPVEFLQEKEVFKEFTYR RNTLGWTSRTQFEETWATLLGVLYTQPLVME QESPPEEDTERTRIVALVAQATISLJSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISLLLLQNPTERGGSM							VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
NITSVTKINADKNAHINHIRLFEPLVIKALKQ YTTTICVQLQKUDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGGPRESEAHIPNIFF FLVLLSYERYHSKQIIGFRIQLCDGIMASGR KAVTHAPALQPVHDLFVLRGTINKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADILPMLAKQQMHI DSHEALGVUNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQGLGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLALR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFRFLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLLAPLSLGMSEISGG(KSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNITPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPL VNSYTRYPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKKEFTYR INTLGWTSRTQFETWATLIGVLVTQPLVME QEESSPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKENNATHHLYQAWD PVPSLSPATTGALISHEKLLLQUNPERPELGSMS	1	·	'				
YTTTTCVLQKQVLDLLAQUQRVHYVCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGFKIIQLCDGIMASGR KAVTHAIPALQPIVI-DLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENBDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSLRPVDMLLRSMF VTPNTMASVSTVQL WISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLQLVGILLLEDIVT KQLKVEMSEQQHTFVQCELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQULLLVNHTDYRWW APVQOTPKRRISLSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSFS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLLAFALSIGMSEISGGKSALFEA AREVTLARVSGTVQQLAVHHVFQPELPAEP AAYWSKLNDLPGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPPLSLDLQAGLDCCCLALQLPGL WSVVSSTEPVTHLEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLAIGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRYPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQKEVFKEFIYR INTLGWTSRTQPEETWATLLGVLVTQPLVMEL QEESSPEEDTERTQINVLAVQAISLVLSAMT VPVAGNPAVSCLEQQPRKPLKALDTTRGRK LSIIRGIVEQEIQAMVSKRENATHHLYQAWD PVPSLSPATTGALISHEKLLLQUPPERELGSMS		İ					
DSDQVFIGFVLKQFEYEVGOGRESEAIPNIFF FLVLLSYERYHSKQIIGPKIIQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTINKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPVLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHITFYCQELGTLLMCLHHEKS GMFRRITAAATRLFRSDGGGGSFYTLDSLINLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHIDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFJQAIQSRCENLSTFINLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLLAANLQSSMQLPMEELNRIQEY LOSSGLAQRHQRLYSLDEFRISTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLACLSGAMSEISGGKSALFEA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHAASSEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISSLALPLVNSYTRYPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWAILLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAJTSLVLSAMT VPVAGNPAVSCLEQOPRIKPLKALDTRGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTQALISHEKLLLQNLYTQFLVME QEESPPEEDTERTQINVLAVQAJTSLVLSAMT VPVAGNPAVSCLEQOPRIKPLKALDTRGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTQALISHEKLLLQNLYTGPLXME	1	Ì .					
FLVILSYERYHSKQIIGPKIIQ.CDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASSYTVQLWISGLAIR.VLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGSFYTLDSLNLR ARSMITHPALVLUWCQILLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHOJLISLSHEPPVODFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLICTPFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDMYVHLVK SQCWTRSDSALLEGABLVNRPAEDMMAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLAIQLPGL WSVVSSTEFYHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRYPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLGLYLVTQPLVME QEESPPEEDTERTQINVLAVQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALSHEKLLLQNDFERELGSMS	1						
KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPVLISCTVINRLRDGDSTSTLE EHSEGKQKNLPEETFSFILQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREFVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSIPTMLKKTLQCLEGI HLSQSGAVITLYVDRLLCTPFRVLARMYDDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFFRLSTIMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLAPCLSLCMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPPAEP AAYWSKLNDLFFGAKLYDVFQPLPAEP AAYWSKLNDLFFDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKKDIVKFVVATLEAL SWHLIBEQPIELDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTFKAISEEBEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEFVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTVPLVWKL GEESPPEEDTERTQINVLAVQAITISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LETOKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIIIPMLAKQQMHI DSHEALGVINTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKINLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLIMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRNVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQGLDCCCLALQLPGL WSVVSSTEFVTHACSLTYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAPPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGVLVTQPLVME QEESPPEEDTERTQNNVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIEGIVEQEIQAMVSKRENIATHHLYQAWD PVYSLSPATTGALISHEKLLLQINPERELGSMS	1						
CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQLLLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISJSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPIMLKKTLQCLEGI HLSQSGAVCTLLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELWRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA ARREVTLARVSGTVQQLPAVHHVPQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLOSVALIGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQAFFTPVFFLVEKKLDUTRPGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVYSLSPATTGALISHEKLLLQINPERELGSMS		1					
DSHEALGVLNTI.FEILAPSSLRPVDMLL.RSMF VTPNTMASVSTVQLWISGILAIRVLISQSTED IVLSRIQELSFSPYLISCTVINRILROGDSTSTILE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIIFKS GMFRRITAAARTHFRSDGCGGSFYTLDSLNLR ARSMTTIHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIFVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLYTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1	!					
IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNJEETTSRFILQLVGILLEDIVT KQLKVEMSEQOHTPYCOELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDBFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVSSKLPSHLHLPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTPKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS			·				
EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIJFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNRETVRGGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1				:		
KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHISLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSITPIMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHILYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	1					
ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHIDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSFILHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHILYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		i					KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHIDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDDL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
HLTWLIVNHIQDLISLSHEPPVQDFIŠAVHRNS  AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLIQINPERELGSMS	1						
AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							HI TWI IVNHIODI ISI SHEPPVODEISAVIDAIS
HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						AASGLEIOAIOSRCENI STPTMLKKTI OCLEGI
ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS							
LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1						ACRRVEMLLAANLQSSMAQLPMEELNRIQEY
PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS							LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	]					PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNIPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		}					AREVILARVSGIVQQLPAVHHVFQPELPAEP
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGDDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNIPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGDDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	]					
EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							WOVVOOTEFVTHACOLIVOVITIER EAVAVORO
TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						EOLL SPERRTNTPK AISPPEPEDDDNTONDEN
FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	}						
INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS						İ	
QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							INTLGWTSRTQFEETWATLLGVLVTOPLVME
VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							QEESPPEEDTERTQINVLAVQAITSLVLSAMT
PVPSLSPATTGALISHEKLLLQINPERELGSMS		· !					VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE	1					ŀ	LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE	1						PVPSLSPATTGALISHEKLLLQINPERELGSMS
	L	<u> </u>					YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE

000 70	Long		T-0550	[ n. 4: 1 ]	I D. 453 1 1 1	LA :
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of,	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	in			I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding	
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	[		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	Ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	!	ļ		peptide		/=possible nucleotide deletion, \=possible
· · · · · ·	ļ			sequence		nucleotide insertion
	ì		ł	ł		ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
	1		1		İ	LELYSRWILPSSSARRTPAILISEVVRSLLVVS
	1 .	i	1			DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
		İ				YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
	1	ļ		<b>!</b>		RSSHLPSRVGALHGILYVLECDLLDDTAKQLI
	[	ľ	ĺ	•	[	PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT
						AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE
			ļ			ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV
					ļ	KLSVDRVNVHSPHRAMAALGLMLTCMYTG
	ļ			1		KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL
	l		1			FDRIRKGFPCEARVVARILPOFLDDFFPPODIM
						NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS
	ŀ			i		TGQSSMVRDWVMLSLSNFTQRAPVAMATWS
	1	•				LSCFFVSASTSPWVAAILPHVISRMGKLEOVD
	i	I	1	ĺ		VNLFCLVATDFYRHQIEEELDRRAFQSVLEV
						VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA
720	1776	Α .	3447	-	130	WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG
	İ	ļ	1	-	ļ	LPCVGDAAEYODCNPQACPVRGAWSCWTS
			}			WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL
						GLHTEEALCATQACPEGWS
429	1779	<del> </del>	3464	583	3	
429	17/9	A	3404	363	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/
	l	ĺ			ĺ	RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR
	l				ļ	AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP
	1				1	CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI
	j				•	PPES/RS*QGGTVQTGQHSSGREAGSWRARGR
			<del>]</del>			NAGRR*KGGGKIGTKQGAVRARKECRGEMA
						SGETDSE
430	1780	A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA
	ļ		i			HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR
			l			VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE
	1	l	1			ASKCNIVCTQPRRISAVSLANRVCDELGCENG
	l		1	Ì		PGGRNSLCGYQIRMESRACESTRLLYCTTGV
						LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS
	1					DFLLIILKEILQKRSDLHLILMSATVDSEKFST
•	]	ļ		ļ	ļ	YFTHCPILRISGRSYPVEVFHLEDIÆETGFVLE
	· .			1		KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE
		l		l		YIPVQTGAHADLNPFYQKYSSRTQHAILYMN
	1	1	ł	1		PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL
			1	1		PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI
	1	ł	ĺ	1	1	LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI
	1			1		PDVVFVIDTGRTKENKYHESSQMSSLVETFVS
	1			1		KASALQRQGRAGRVRDGFCFRMYTRERFEG
	l			1	}	FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF
		J	]	1	l .	LSKALDPPQLQVISNAMNLLRKIGACELNEPK
		1				LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP
						VATLAAVMTEKSPFTTPIGRKDEADLAKSAL
				1		AMADSDHLTIYNAYLGWKKARQEGGYRSEI
		1		I	İ	TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF
	l	l	ł	1	l	SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA
	1	1	-		1	GLYDNVGKIIYTKSVDVTEKLACIVETAQGK
	1	1	1	1		AQVHPSSVNRDLQTHGWLLYQEKIRYARVY
		1	[	1		LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW
						IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK
	1		[	1		
			1	1	1	MSLENDKILQIITELIKTENN
401		<del></del>		<del>  .                                     </del>		
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ
431	1781	A	3474	1	441	PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL
431	1781	A	3474	1	441	PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA
431	1781	A	3474	1	441	PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL

ero ID	SEO ID	Met	LCEO	Desdisted	Predicted and	I Amino soid so (4 - 4) i - C-C-4
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	<b>!</b> ,			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	`		1	peptide	•	/=possible nucleotide deletion, \=possible
			1	sequence		nucleotide insertion
432	1782	A	3478	416	23	QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
		ł			<u> </u>	QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
		l				CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS
	ĺ	ĺ	[	ĺ		QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY
					į	SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS
						SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL
	l		ł			QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS
						APRSRCVARPAARTGLPTPAPASSPAPAASPA
						PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
						GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
		1	· ·			AASPSPAASPAPPAASPVLTASPPLPAASPSPA
	ĺ	1	1	1	ĺ	ASPAPPAASPVLTASPPLPAASPALAASPVHT
•	1		1			ASPPVHVASPPVHTASPPVHVASPPVHTASPP
						VHVASPPVHTASPHVHVASPPVHVASPPVHV
	1	i	1		i	ASPPVHTASPPVHVASPPVHTASPHVHVASPP
	ļ					VHTASPPVHVASPPVHVASPPVHVAYPPVHV
						ASPPVHVASPPVHVASPPVSCSGDSTSDCFPP
		ļ	1			QPGAVFPHSLAPSLGGWSHLVAALP
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
			1			SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ
	ļ	i	1			SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ
			1		i .	WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV
435	1785		3529		2161	RTKFGINMVTSRERGTTRLPKEG
433	1/65	A	3329	1	3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA
		ŀ				DEVQHCQSILNSLLPRASTSKEVDASLLSVVS
		İ				FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF
	İ					LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC
		]	ļ ·			EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYONPHT
		}				VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
		l				PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
	ŀ		j			SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL
						KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF
						LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK
		Ì				LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE
		<b> </b>				MLRTDLSYLCSRWRMTGQPTTTFPISHSMLDE
		f				DGTSLNSSILAALRKMQDGYFGGARVQTGKL
						SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN
	Ì	İ				YDYLESGNWMNDYDSTSHARCGDEVARYL
		<b>!</b>				DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT
		}	}			TCDLMSLVTKAKELHVQNVHMYLPTKLFQA
		1				SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ
						SGEVDFKALVLQLKETSSLQEQADILYMLYT
		1	'			MKGPDWNTELYNERSATVRELLTELYGKVG
		1				EIRHWGLIRYISGILRKKVEALDEACTDLLSH
		1		ı	1	QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA
						SEGDMSISILTQEIMVYLAMYMRTQPGLFAE
		ŀ				MFRLRIGLIIQVMATELAHSLRCSAEEATEGL
		l				MNLSPSAMKNLLHHILSGKEFGVERSVRPTD
		٠ .				SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK
						QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
		1				QRRRRLDGALNRVPVGFYQKVWKVLQKCH
1		l.				GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL
				,	[ ·	NRVPQPEYRQLLVEAIL\VLTMLADIENHSIGS
			<b>j</b> 1	•	1	THE TREE PROPERTY OF THE PROPE
						IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD
				•		IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY
				·		IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ
436	1786	A	3546	73	393	IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	delice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		]	1	peptide	Ì	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
437	1787	A	3554	5157	2939	PXSARSCWMRKG AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG
437	1707	<b>^</b>	3334	, ,	2,55,	SASGCRYAAHPHWGLGGAAAAGGSWEPQPP
				,		RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA
			ļ			AQPGARARTSPPPASARNMAARPAATLAWSL
		Ì				LLLSSALLREGCRARFVAERDSEDDGEEPVVF
			Ì		ļ	PESPLQSPTVLVAVLARNAAHTLPHFLGCLER
		[	1			LDYPKSRMAIWAATDHNVDNITEIFREWLK
						NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV
			1			DNFLTNPOTLNLLIAENKTIVAPMLESRGLYS
						NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP
		1	1			VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW
			1			TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP
		1	ĺ		ĺ	LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ
		1				YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD
1				!		RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA   LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV
		1				WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK
						LMDNIDQAQLDWELIYIGRKRMQVKEPEKA
						VPNVANLVEADYSYWTLGYVISLEGAQKLV
'			ĺ	1		GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
			İ			YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE
		}			1	TSTIWDNETVATDWDRTHAWKSRKQSRIYSN
438	1788	A	3563	130	527	AKNTEALPPPTSLDTVPSRDEL IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C
130	1700	Α	3303	150	327	NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA
		ŀ				NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA
1						SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT
						TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLLSTLTGRSY
						GQPSLQDELKDNTTVFTRILDRLLDGYDNRL
		ļ	}			RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI   DVFFRQSWKDERLKFKGPMTVLRLNNLMAS
				,		KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE
						DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH
1		ł	1			ACPLKFGSYAYTRAEVVYEWTREPARSVVV
1						AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV
			1			MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF
		1	}			WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN
		[				YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN
			]			NTYAPTATSYTPNLARGDPGLATIAKSATIEP
		1	1			KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
L			j		_	FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK
[			[			FWEVISDEHGIDPTGTYHGDSDLQLERINVYY
						NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG
441	1791	A	3569	2	1751	GKYVPRAVLVDMEPGTMDSV   FVAVAGAVSGEPLVHWCTOOLRKTFGLDVS
441	1/91		2000	[ *	1731	EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGO
1						FIEELITKWQKNDQELISDPLQQCFKKDEILDG
ł		ł	1	}		QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE
	-	l				VKTPFDLAKAQENSNSVKKKTKFVNLYTREG
			j	]		QDRLAVLLPGRHPCDCLGQKHKLINNCLICG
						RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS
			Ì			NKSQKLLKKLMSGVENSGKVDISTKDLLPH
L	L	<u> </u>	L	L	L	QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first amino acid		
					of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	į			peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
						DDESDYFASDSNQWLSKLERETLQKREEELR
						ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
	İ					SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
		]				VNPNMYQSPPQWVDHTGAASQKKAFRSSGF
						GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH
		1	}			QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
			1	1		TAKKPSPQEVSELQATYRLLRGKDVEFPNDY
						PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS
			ļ			PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
						QGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK
						KTAGVKPYECTICGKAFMRLSSLTRHMRSHT
	·		1			AIRAI\EKPYKCKEC\GRAFSLSQILSK\HERSH
			l			TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK
	1	1				PYECKQCGKALSCSSSLRVHERIHTGEKPYEC
	1	1	1	1		KQCGKAFSCSSSIRVHERTHTGEKPYACK\EC
		1				GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
	l	1				FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
	1			1		TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV
		1	1		•	HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
<b> </b>	l	1				HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG
		1				EKPYECKECAKTFISLENFRRHMITHTGDGPY
						KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
	l	ľ				CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK
						AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
						LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
İ		1				LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
		1				QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
		!		1		PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL
i		1				VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
			ŀ			DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
				<u></u>		ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK
	<u> </u>	<u> </u>			L	LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
	1	ľ		1	1	MTKVTLENFYSNLIAQHEEREMRQKKLEKV
		1.			1	MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
					1	RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH
				-	1	VYAMKILRKADMLEKEQVGHIRAERDILVEA
	1	1			1	DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM
		1			1	MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL
1		1			•	GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
					1	KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE
					1	TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
]					1	KLCDWWSLGVIMYEMLIGYPPFCSETPQETY
1					1	KKVMNWKETLTFPPEVPISEKAKDLILRFCCE
1	ł	1	1		1	WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
1	1	1			1	AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE
1	1				I	TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
					1	KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD
ļ	'	_	1			GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK
1				1	1	NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
	1	1		l	1	FCKLGVLVRHRISLFGNDATSIVNCLHILGOT
				1	1	LDARTVMKTGLESVKSALRAFLDNAAEDLE
ŀ	1					KTMENLKQGQFTHTRNQPKGVTQIINYTTVA
			ı	1		I I DI U COI ECITICONOCCEDI A EDITOROLOGIA
	i.	1	l .	1		LLPMLSSLFERIGOROFGEDLIFEDVEDVSCVV
						LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERORSALGECLAAFAGA
						LTSLYALGTSKSIYVERQRSALGECLAAFAGA

SAU ID NO. of pepide of the period of pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of the pepide of pepide of the pepide of pepide of the pepide of pepide of pepide of the pepide of pepide	SEO ID	SECTO	Mat	SEC	Dradioted	Dredisted and	Amino acid sequence (A=Alanine C=Cysteine,
nuclectide seq- uence ue	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
uence  USSN postor uence  0944  14  15  USSN postor			DOG	_		· ·	D-Aspartic Acid, E-Giutantic Acid,
uence  914  914  916  917  918  918  919  919  919  919  919							
uencc    914   ng o first amino acid residue of peptide peptid	eotide	seq-					
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per solo podon, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide insertion    MPHYMEVILEPALCSYMSRW WEHGPENPER   AEMCCTALNSHMNTLLONILKINYNLGIDE   GAWMERLAVFSOPINKVERQLEXTHEFLEM   EKILKKAATVVSEDHLAKAARAGOMSERGE   LILDEFITLARDLYAAFYPLLIRFODYNRAKWL   KEPNPAELEL FRAMVAEVPINYSKSHMYRKER   ENKINKKAKATVVSEDHLAKARAGOMSERGE   QNFVVQNENNMSFLITDTKSKMSKAAVSDQ   ERKKMEKGDHYSMQTSLVAALKALPLEM   ERKKMEKARDHYSMGTSLVAALKALPLEM   NIHLOGALEDPAIRWQMALYNDLPRETDER   NIHLOGALEDPAIRWQMALYNDLPRETDER   DPEKTVERVLDIANVLFHLEQKSKRVGRRHY   CLEHPROSEKKAWHKLLSKQKRRAVVAD   CLEHPROSEKKAWHKLLSKQKRRAVVAD   ERKKMEKAL EDDELTMAAT ADIMAKS   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   UTENSGEVLODDEFITCHEDETSTRYDLAK   EKINGLIANGONSTVQCKMLDPKEKKAWH   LILLISRITALTEKCKLEDDELTMAA ADIMAKS   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDOSEERVKSFEERBESTCKRVDHAY   EKINGLIANGONSTVQCKMLDPKEKKAWH   VTERSGEVLADDEFITCHSFLOLLED   NSDPQNYLRTQTGNNTTVNIIISTVDLLRVQ   ESISDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW      ESISDFWYSGKDULBOGORNE	seq-	uence		09/496			M=Methionine, N=Asparagine, P=Proline,
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per solo podon, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide deletion, /~possible nucleotide insertion    MPHYMEVILEPALCSYMSRW WEHGPENPER   AEMCCTALNSHMNTLLONILKINYNLGIDE   GAWMERLAVFSOPINKVERQLEXTHEFLEM   EKILKKAATVVSEDHLAKAARAGOMSERGE   LILDEFITLARDLYAAFYPLLIRFODYNRAKWL   KEPNPAELEL FRAMVAEVPINYSKSHMYRKER   ENKINKKAKATVVSEDHLAKARAGOMSERGE   QNFVVQNENNMSFLITDTKSKMSKAAVSDQ   ERKKMEKGDHYSMQTSLVAALKALPLEM   ERKKMEKARDHYSMGTSLVAALKALPLEM   NIHLOGALEDPAIRWQMALYNDLPRETDER   NIHLOGALEDPAIRWQMALYNDLPRETDER   DPEKTVERVLDIANVLFHLEQKSKRVGRRHY   CLEHPROSEKKAWHKLLSKQKRRAVVAD   CLEHPROSEKKAWHKLLSKQKRRAVVAD   ERKKMEKAL EDDELTMAAT ADIMAKS   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   UTENSGEVLODDEFITCHEDETSTRYDLAK   EKINGLIANGONSTVQCKMLDPKEKKAWH   LILLISRITALTEKCKLEDDELTMAA ADIMAKS   CHEEDDDOSEERVKSFEERBESTCKRVDHAK   CHEEDDOSEERVKSFEERBESTCKRVDHAY   EKINGLIANGONSTVQCKMLDPKEKKAWH   VTERSGEVLADDEFITCHSFLOLLED   NSDPQNYLRTQTGNNTTVNIIISTVDLLRVQ   ESISDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW   VTERSGEVLADDEFITCHSFLOLLED   SSINDFWYSGKDULBOGORNESKALDW   ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW    ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW     ESISDFWYSGKDULBOGORNESKALDW      ESISDFWYSGKDULBOGORNE	uence			914	ng to first	acid residue	
residue of peptide sequence     Y=Tyrosine, X=Dukanown, *=Siop codon, /-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide delicin, \*-possible nucleotide insertion     APM-WEWILPMC.SYMSR.WEHGPENNFER AEMCCTALNSEHMITLLGNILKIIYNNLGIDE GAWMRKILAYSOPIDIK VEPOLLKTHFLPLM EKLKKKAATVVSEDHILKAARAGIMSEAEL LIDEFTLARDLYAFPLIIRFODYNRAKWL KEPNPEAELFRMYAEVITIVWSKSINFKREE ONFVO/DREINMSELTITUTISKMSKAAVSDQ ERKKMKKGRYSMOTSLIVAALKRLLPIGL NICAPGOEQLIALAKNRTSIKDTEDEVRDIRS NIHLQGKLEDPAIRWOMALYKOLPNRTDDTS DPEKTVERVIDLANVILLGOYEKSWGRHY CLVEHPORSKKAVWHKLLSKORKRAVVACF RAMPLYNLFRHRAVNLLGOYEKSWGRTHY CLVEHPORSKKAVWHKLLSKORKRAVVACF RAMPLYNLFRHRAVNLLGOYEKSWGRTHY CLVEHPORSKKAVWHKLLSKORKRAVVACH REDISTATATEKCHEEDFILMAYADHAKS CHDEEDDDGEEVKSFERKEMEKOKLLYQQ ARLIEDRAAFBOAAEMVLQTIASSKGETDPWAAT LKLGIAILNGGNSTVQOKMLDYLKEKKDVGF POSKAGLMOSCSVLDLAFREQNKASGLIM VTEEGSGEKVLQDDEFTCDLFRLQLLCEGH NSDEQNYLRTGTGNNTLVMISITSVDYLLRVQ ESISPFWYYSGKOVIDEQGGRNESKAIQVA KQVPNTLTTYVGTSAKGETDFWAAT LKLGIAILNGGNSTVQOKMLDYLLEKKDVGF POSKAGVMALLSMLEGNYNVATIGTISCKGWOMML VEGSSGEKVLQDDEFTCDLFRLQLLCEGH NSDEQNYLRTGTGNNTLVMISITSVDYLLRVQ ESISPFWYYSGKVUDEQGGRNESKAIQVA VQFLHVFAHMQMKLSQDSSQELLKELMDLQ KOMPANALSMLEGNYNVATIGTISCKGWOMML VGFLHVFAHMQMKLSQDSSQELLKELMDLQ KOMPANALSMLEGNYNVATIGTISCKGWOMML VGFLHVFAHMQMKLSQDSSQELLKELMDLQ KOMPANALSMLEGNYNVATIGTISCKGWOMML VGFLHVFAHMQMKLSQDSSQUELKELMDLQ KOMPANALSMLEGNYNVATIGTISCKGWOMML VGFLHVFAHMQMKLSQUESDSLAMBANAKESSEKRPPE QFRAMFSILTVRSALFARNYILTIMMLS LKSLKKQMKKYKMTVKDMVTAFFSSYWSI POSKGVKKFFTOWNTGGGGKEKEMELFYNFCED TIFEMQLAAQSGLINGTSENLAMATHA VGFLHVAYSFRGFFTGLSLLIGGSCVEGA KKKVASLALARNYILTIMMLS LKSLKKQMKKYKMTVKDMVTAFFSSYWSI PMTLLHFVASVFRGFFTGLSLLIGGSTUKMTYTHPSUFF PANAWARTSILTVRSALFARNYILTIMMLS LKSLKKQMKKYKMTVKDMVTAFFSSYWSI PMTLLHFVASVFRGFFTGLSLLIGGSTUKMTYTHPSUFF PNTWYMFYKMTARNYILTSSYTH NICKQLVTTVAFFRKF LEEGGGGYKLMALTYTVAFFFTKF AKSENDA KYTSLDSSHRIJAHVYLGESSGYMPTVRL SYNTH NICKQLVTTVAFFFTKF AKSENDA K					amino acid	of peptide	
peptide sequence    Poptide	I '						
sequence    Sequence   New York Purple   New Yor	1			j .			
MPHYMEVILPMILGSYMSRWEHOPENNER ABMCTALNISHMINTLIGHILKITYNILGIDE GAWMRILAVFSQPINKVKPQLIKTHFLPLM BELKKKAATVVSEGHIKAKPAGGISKEAEL LIJ.DEFTTLARDI, VAFYPLLIRFGDYNRAKWL, KENPPEAELERMYAEVFIVWSKSHNFKREE QNFVVQNEINNMSFLITDITKSKMSKAAVSDQ BEKKMKRIGDRYSMGALIVELPIGI NICAPGDQELLALAKNRFSI, KDTEDEVRDITRS NIHLQKGLEDPARIWGMALI YADLPHSTIDITS DPEKTVERVLDIANVL, FILEGKSKR VQRRHYY CLYEMPORSKAAVWKLISKQRKRAVVACF RMAPLYNLPRIBAVNLLFQUSKSKR VQRRHYY CLYEMPORSKAAVWKLISKQRKRAVVACF RMAPLYNLPRIBAVNLLFQUSKSKR VQRRHYY VEIKLIBLIAKPG ABPEEDIGGTKFVDPLHQ LILLISRTALTEKCK LEEDFLWAAYADIMAKS CHDEEDDOGEEVKSPEREMENKOLIVAQ ABLIDRGAABMVLQTISASKGETOPMVAAT LILGIALINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKALILINGGNSTVOJCKALDYLKEKKDVOP POSKAVINAVALINGGNSTOJCKALDYLKEKALOV ROMVVMLISHLEDNIV VNOTIGKCOMILLEGH NEDDOVILETOTGNNTTVNILISTVOYLLEVOL ESISPPVAVYSCKOVKALDYLKEKALOV KQYPNTLTETYQOPCTCHONQSLAHSRLWDDAV VGFHHYARMOKLSQDSSOJELKELINDLQ KDMVVMLISHLEDNIV VNOTIGKCOMILLEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH SSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH VKSSROVEMILISTVOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSTOVATATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROVEMILITATION TOTTOJCHALEGH VKSROV	1	[			- •		
AEMCCTALNSEMMITLIGNILKINYNLGIDE GAWMRIAVTSOPIINKVEPQLIKTEPIPAM EKLKKAATVVSEEDHILKAARAGIMSIAEL JLDEFTTARDI VAFPULIRFODYNRAKWI, KEPNPEAEELERMYAEVFTVWSKSHNFKREE ONFVOWENDMISSITTDTKSMSKAAAVSDO BEKKMKRKGRYSMOTSLIVAALKRLLPIGI NICAPODOELLIALAKNETSLKDTEEDVEVDIURS NIHLQGKLEDPARWOMALYKDLENRTDDTS DPEKTVERVIDLANVLFHLEQKSKRVGRHY CLVEHPQRSKAAVWHKLLSKQRKRAVVACF RMAPLYNLPRIRAVNLFLQGYEKSWETEEH YFEDKLEDLAKPGAEPFEBDGTRKTVDPHQ LILLISSTALTSKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEVEKSPEKEMEKKKUTVDPHQ LILLISSTALTSKCKLEEDFLYMAYADIMAKS CHDEEDDOGEEVEKSPEKEMEKKKUTVDPHQ LILLISSTALTISKCKLEEDFLYMAYADIMAKS CHDEEDDOGEEVEKSPEKEMEKKKUTVDPHQ LILLISSTALTISKCKLEEDFLYMAYADIMAKS CHDEEDDOGEEVEKSPEKEMEKKKUTVDPHQ LILLISSTALTISKCKLEEDFLYMAYADIMAKS CHDEEDDOGEEVEKSPEKEMEKKKUTVDPHQ LILLISSTALTISGASCAVLDLAERGNKAEGLGM VTEGSGEKVLQDDEFTCLIFERLJLLEGH NSDPQNYLRITGIGNITTVNIIISTVDYLLRVQ ESISPFYWYSGKOUPEGOGRAFSKAIQVA KQVFNITLTEYIGOPCTGNQOSLAHSRUMDAV VOFLHYVATMOMKLOSSOSIELLKEMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNOVEMEMKFPDMFLKLKDLTSSDTYFKEYD PDGGGVFKRDPHKAMESHKHYTQSETFELL SCARTDENTLDVEEFEVRFHEARDIGFNVA VULTNI.SEHMPNDTRLQTIFLELAESVUNYPCPP FLGREIMGSAKRERVYFEIESSSRYMEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED  TIFEMGLAAQIESGLIANSPYPTGWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED  UKSKKKVFKVKMVMVTLAFSSYWSI LKSLKKOMKVKKKMVKMVMVTAFSSYWSI PMTILHFVASVSVGFKEPTRSSSNA KKIVALAAQIESLLAGSLLSDIFGIDLIKE EGGGYKLIPHTVAGLEDLASSICLLIGGSLVEGA KKIKVALAALISSTALLIGSSLVEGA KKIKVALAALISSTALLIGSSLVEGA KKIKVALALAMPOPTODEVRGGGEGEERKP LEAALPSEDLTDLKELTESSDLLSDIFGIDLIKE EGGGYKLIPHTVAGLSDLMSNPYPVURG KFGROKAKEEKEEKEERSSSPEKAEGEDGE KEEKAKRDKGKOKK,RKNVKDMVTAFSSYWSI PMTILHFVASVSVOGKELPTRSSSNA KKIVALDASSHRILATHYVLEESSGYMEPTVRIL PLHTVISPECIGIGYCLKPLTTRSSSNA KKIVALDASSHRILATHYVLEESSGYMEPTVRIL PLHTVISPECIGIGYCLKPLTTRSSSNA KKIVALDAGGEREER LEPGG.YTTEOPSEDDIKGOWDRLVNTONSP NAYWDKFFRAHLDLIANGFRILTRITISSVTH NGKQLVLTVGLLAVVVVTLYVAFRFRKF YKKEIDOTOPSEDDIKGOWDRLVTNTINSVTH NGKQLVLTVGLAMPKVGERVEEREWETKSFFFFF TORGIGDBEFFFFFTHANANNLTLSVTH NGKQLVLTVGLAMPKVGERVEEREWETKSFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	<b></b>	<u> </u>			Sequence	<del> </del>	
GAWMKRIAVFSQPINKVRPQLIKTHEPLIN ELKKKAATVVSEEDHIA.FAGEMSEAEL ILIDEFTTI.ARDI.YAFYPLILIRGGDYRBAKU. KEPRPFAEELFRWAVEVPIYWSKSSINFKREE QNFVVQNEINNMSFLITDTKSKMSKAAVSD ERKKMARKGBRYSMQTSLIVAALRILIPIGL NICAPGDQELIALAKNEFSIKDTEDEVRDIRES NIHLQGKLEDPAIRGWQMALYKDLPNRTDDTS DPEKTVERVLDIANVLFHLEQKSKRVGRAVVACF RMAPLYNLPHRAVNLFLGGYEKSWETEEH YEDKLIEDLAKPGAEPPEEDEGTKVPDHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDEDDDGEEVKSFEKEKMEKPUPHQ LILIFSRTALTEKCKLEEDFLYMAYADDMAK CHDENDDGEFTCDLFRTWDPHA LILIGALINGGNSVOLDEGGGREGOMEN VEEGSGEKVLQDDEGTCDLFRCKULTEKDULCEGH NSDPONYLRTOTONNTTONIISTVDYLLRVO ESIDDFWYYSKGKUNDEGGGGREGAKADVA VOGHLEVFAHMOMKLSODSSOGLILKELMDLQ KDMYVMLISMLEGNVYNOTIGKGMVDML ESSNNVEMLKFFDMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSNNVEMLKFFDMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSNNVEMLKFFDMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSNNVEMLKFFDMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSNNVEMLKFFDMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSNRVEMLKFFNMFKLKLDLSSDTFKEYD PDGKGVFFKDFHKAMESHKHYTOSSTEFLY ESSRCPHTDVNRGGGEKKMELFVNPCED TFEMPQLAGNSESDLNESANKERSTVOPEKPQ FUGREINGSAKREVYFEISSTTONEKPQ FUGREINGSAKREVYFEISSTTONEKPQ FUGREINGSAKREVYFEISSTTONEKPQ FUGREINGSAKREVYFEISSTTONEKPQ FUGREINGSAKREVYFEISSTONEKPQ FUGREINGSAKREVYFEISSTONEKPQ FURDELALANMPDFLQDEVRGDGEGGREKME SILKERKOKKMATULDATTAFFSSYWSI FMTLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDFLQDEVRGDGEGGREKME SILKERKOKKMATULDATTAFFSSYWSI FWKLIKLTYNVYTLESSOFTMETTVRT FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQUFTRSSSWYSI FURDELALANMPDFLQDEVRGDGEGGREKME FURDELALANMPDFLQUFTRSSSWYSI FURDELALANMPDFLQUFTRSSSWYSI FURDELALANMPDFLQUFTRSSSWYSI	1						
EKLKKKAATVVŠEEDHLARGROMSRAW    LILDEPTTIARDI, YAFYPLIRGROFONRAKW    KEPNPEAEELFRMWAEVFIYWSKSHINKRISE    QNFWVONDENNINSHITDTINSKISKAAVSDQ   ERKKMKRKGDR YSMQTSLIVAALKRIL FIGH   NICAPGOGLIALAKANFSLKDTEDEVRDIIRS    NIHLIQGKLEDPARRWQMALYKDLPNRTDDTS    DPEXTYDERVLDIANVIHLEQKSKYGERHY    CLVEHPORSKKAWHKLLSKORKRAVVACC    RNAPLYNLPRHRAVNLFLOGOYEKSWIETEEH     YPEDKLEDLAKFOAEPPEDEGTKRVDPLHQ    LILLFSRTALTEKCKLEEDFLYMAYADINAKS    CRIDEDDGEEEVKSPEREMEREKCKLLVQQ    ARLHDRGAAEMVLQTISASKGETGFMVAAT    LELGGIALINGGNSTYOQKMLDYLKEKDVGF    FOSLAGLMIQGCSVLDLNAFBRONKABGLOM    VTEBGSGEKVLQDDEFTCLLFRTQLLCEGH     NSDPQNYLRTQTGNNTTVNIISTVDYLLRVQV    ESISDFYWYYSKGNUDEQGQRIFKALGUAGN     VTEBGSGEKVLQDDEFTCLLFRTQLLCEGH     NSDPQNYLRTQTGNNTTVNIISTVDYLLRVQV    ESISDFYWYYSKGNUDEQGQRIFKALGUAGN     VOFLHVAHMQMKLSQDSSQELLKELMDLQ     KOMVVALLSMLEGRVVNOTIGKQMVDMLV     ESISNFWYNYSKERDFLKLKDLTSSDTFKEYD     PPGKOVFRKDFHKAMESHKHYTQSETEELL     SCAETGENETLOYEFVKRTPHFAKDGFNVA     VLLTINLSEHMPNDTRLQTTLELAESVLNYPOP     FLGREIMGSAKREVVYFEIESSRTWKEKPV     PPGKOVFRKDFHKAMESHKHYTQSETEELL     SCAETGENETLOYEFVKRTPHFAKDGFNVA     VLLTINLSEHMPNDTRLQTTLELAESVLNYPOP     FLGREIMGSAKREVVYFEIESSRTWKEKPV     PRGKOVFRKDFHKAMESHKHYTQSETEEL     SCAETGENETLOYEFVKRTPHFAKDGFNVA     VLLTINLSEHMPNDTRLQTTLELAESVLNYPOP     FLGREIMGSAKREVVTFEIESSRTWKEKPV     VKESKROFIFDVNEGGEKKEMELFVNRCED     TIFMQLAQGISEDLNERANKESSEKERPE     QGPRNAFFSILTVRSALFALRYNILTLIMMUL     LKLKKKKMKVKKMTVMDMYTAFFSSYWSI     FMTLHFVASVFRGFFRICSLLLGGSLUGG     KPQGLAAGEEKEEEEETKSEPFKAEGGEGGE     KPQGLAAGEEKKEEKEETKSEPFKAEGGEGEK     LEAALPSEDLTDLKELTEESDLLSDFFGDLKE     KPGGCKAKEEKEEKEETKSEPFKAEGGEGEK     LEAALPSEDLTDLKELTEESDLLSDFFGDLKE     ARMTHLLFYKVALFALRYNDYNMRMLALFU     PHTVISFFCIICIYYCLLVPL VITEKEEVARK     GGGYKLPHNINGAGLSDLKNSNYPWPEVQE     KPGGCKAKEEKEEKEETKSEPFKAEGGEGKE     KPGGCKAKEEKEEKEETKSEPFKAEGGEGKE     KPGGCKAKEEKEEKEETKSEPFKAEGGEGKE     KPGGCKAKEEKEEKEETKSEPFW     ARMTHLLFT     NICKGLUNTYARNYNNRMLALTS     NICKGLUNTYARNYNNRMLAUR     NICKGLUNTYARNYNNRMLAUR     NICKGLUNTYGENE     NICKGLUNTYGENE     NIC	1						
LILDEFTILARDLYAFYPLLIREGOYNRAKUNE   KEPPPEAEELFRWAEVFIYWSKSINFKREE   QNFVVQNEINNMSFLITDTYSKMSKAAVSIG   DEKKNARKORDS SWOTSLIVAALRIL PIGI   NICAPGDOGLIALAKNEFSLKDTEDEVRDIURS   NIHLQGKLEDPARWQMALYAFUNFTDDTS   DPEKTVERVLDIAVLFHLEQKSKRVGRRHY   CLVEHPQRSKKAVMHKLLISKQRKRAVVACC    RNAPLYNLPHRAVNLFLQGVEKSWETEEH   YFEDKLIEDLAKFOAFPEEDEGTKRVDPHLOW   LILLISRTALTEKCKLEEDFLYMYADIMAKS   CHDEEDDDJGEEWKSFEEKEMEKQKLLYQQ   ARLHPRGAAEMVLQTISASKGETWAVATI   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   FOSLAGLMQSCSVLDLNAFERGMAVAAT   LKLGJAILNGONSTVQQKMLDYLKEKKDVGF   POSLAGLMQSCSVLDLNAFGRANGLANGLANGLANGLANGLANGLANGLANGLANGLANGL	[						, , ,
KEPPPEAELIFRMVAEVPITWSKSINTKYRES QNFVVONDENNINSHITDTYSKIKSKAAVSDQ ERKXMKRKGDRYSMQTSLIVAALKRILIGII NICAPGOQELIALAKNEPSLIDTEDEVRDIURS NIHLQGKLEDPARWQMALYKDLPNRTDDTS DPEKTVERVLDIANVLPHLEQKSKRVGREPHY CLVEHPQRSKKAVWHILLISQKSRKAVGAPH CLVEHPQRSKKAVWHILLISQKSRKAVGAPH VERVENAVHLLISQKSRKAVVACE RMAPLYNLPRHRAVNLFLQGYEKSWEITERH YFEDKLIEDLAKPGAEPPEEDEGIRKVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGMVAAT LRLGIALINGGNSTVQOKMLDYLKEKKDVGF FOSLAGLMOSCSVLDLNAFERONKABGLGM VTEGGSGEKVLQDEPTCDLFRTQLLCEGH NSDFONYLRTQOKMDTVLKEKKDVGF FOSLAGLMOSCSVLDLNAFERONKABGLGM VTEGGSGEKVLQDEPTCDLFRTQLLCEGH NSDFONYLRTQOKMDTVLKEKKDVGF FOSLAGLMOSCSVLDLNAFERONKABGLGM VTEGGSGEKVLQDEPTCDLFRTQLLCEGH NSDFONYLRTQOKMDTVLGEHKENDVG EISDPVMYSKGKUDDEQGGRISKALQVA KQYPHILTEVIGQFCTGNQSLAHSRLWDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGPLHVFAHMQMKLSQDSSOGELIKELMDDA VGRAFGELITVORGEGKEKMELFUNFCED TIFEMQLAAQISESDLNERSANKEESEKERFYEE QGPRMAFFSILTVAGGGKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERFYEE QGPRMAFFSILTVAGGGKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERFYEE QGPRMAFFSILTVAGGGKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERFYEE QGPRMAFFSILTYSSUFGGREELFGRISSENA KKIKVAELLANMPDPTQDEVRGGDEGEGERE KEKAKEDKGKOKKLVKHTVKDMVTAFFSSYNB PMTLHFVASVFGFFRICSLLIGGSLVEGOLILKR EGGQYKLIPNIPNAGLSDLMSNNPYMPBVQE KYQGOKAKEEKERKEETENSESPEGTORGEFICE KERKAKEDKGKOKKLRQLHTHRYGEPEVPBSA FWKKILAYQGKLLNYFARNFYNNRMALLFY AFANNFILLFYKVYTERSSGYNBETYRT LEFGGLYTTEQPSEDDIKGQWDRJVINTOSPIL PHLHTVISFFCIIGYYCLKYPLIVFREKEVARK LEFGGLYTTEQPSEDDIKGQWDRJVINTOSPIL AFANTERIELFYKVYTONSPILTALWTMIMISVL GHYNNFFFAAHLLDIAMGFKTLRTILSSTHI VXYQMWALGGVYTONSPILTALWTMIMISVL GHYNNFFFAAHLLDIAMGFKTLRTILSSTHI VXYQMWAL	1		}			}	
QNFVVQNEINNMSFLITDTXSKMSKAAVSID BEKKMERGOBP SWQTSLIVAALERLIPIGE NICAPGDQELIALAKNERSIKDTEDEVRDIES NIHLQGKLEDPARWQMALYKDLPNETDDTS DPEKTVERVLDIANVLPHLEQKSKEVGEPHY CLVEHPORSKCA VWHKLLSKOKERAVVACE RMAPLYNLPRHRAVNLFLQCYEKSWIETEEH YFEDKLIEDLAKPGAEPPEDEGTKRVDPHHY CLVEHPORSKCA VWHKLLSKOKERAVVACE RMAPLYNLPRHRAVNLFLQCYEKSWIETEEH YFEDKLIEDLAKPGAEPPEDEGTKRVDPHHY CLVEHPORSKCA VWHKLLSKOKERAVVACE RMAPLYNLPRHRAVNLFLQCYEKSWIETEEH YFEDKLIEDLAKPGAEPPEDEGTKRVDPHHY CLVEHPORSKCA VWHKLLSKOKERAVVACE RMAPLYNLPRHRAVNLFLQCYEKSWIETEEH YFEDKLIEDLAKPGAEPPEDEGTKRVDPHHY LILLISERTALTEKCKLEEDFLYMAY ADIMAKS CHDEEDDDGEEEVKSFEEKBMEKQKLLYQQ ARLHENGAAEMVQTISASKGETGWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKLAGLMGSCSVUDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKSAGLMGSCSVUDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKSAGLMGSCSVUDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKSAGLMGSCSVUDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKSAGLMGSCSVULDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDG FOKSAGLMGSCSVULDLAPERGRWAVAT LKLGJALINGGNSTVQCKMLDYLKEKKDVDA VEESKEVLQDDEFTCDLFRFQLCCHOH NSDROYNLRGGNSKMLGGGNSKMLDWAMDLL ESISDFYWYYSCKDVIDLOGGRKKMCMYDMLL ESISDFYWYYSCKDVIDLOGGRKKMCMYDMILL ESISDFYWYSCKDVIDLOGGRKKMCMYDMILL ESISDFYWYSCKDVIDLOGGRKKMCMYDMILL ESISDFYWYSCKDVIDLOGGRKKMCMYDMILL ESISDFYWYSCKDVIDLOGGRKKMCMYDMILL ESISDFYWYSCKDVIDLOGGRKKMCMYDMYDLL ESISDFYWYSCKDVIDLOGGRKMCMYDMYDLL ESISDFYWYSCKDVIDLOGGRKMCMYDMYDLL ESISDFYWYSCKDVIDLOGGRKMCMAPMILL SCAETDERGTCTOGGRKTMCMAPMAPMILL SCAETDERGTCTOGGRKTMCMAPMAPMILL SCAETDERGTCTOGGRKTMCMAPMAPHILL SCAETDERGTCTOGGRKTMCMAPMAPHILL ESISDFYWYSCKDVIDLOGGRKKMCMAPMAPMILL ESISDFYWYSCKDMCTHAPMSVC AKKEVARAGEREKEERAGETKSEPEKAEGGEGE KEKKAKDKOKGKKMCMATHAPMSVC AKKEVARABACARACKARACKOKGRKATARMILL ESISDFYWTMCMAPMAPHILL ENTITLIFYCHOPSEDDIKGGWDNLVINTOSSELLA KKEVARABACARACKARACKOKGRATHAPMSVC AKANTELIFYKVACKGRCKETEPPINT WYMNFFFAAHLLDIAMGKTLRTILSSVTH NGKOLVILVUSLIVINTOSTLYLAWMTMMSVL GHYNNFFFAAHLLDIAMGKTLRTILSSVTH NGKOLVILVUSLIVINTOSTLYLAWMTMMSVL GHYNNFFFAAHLLDIAMGKTLRTILSSVTH NGKOLVILVUSLIVINTOSTLYLAWMTMMSVL GHYNNFFFAAHLLDIAMGKTLRTILSSVTH		j .					LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL
ERKKMERGDRYSMOTSLIVAALKRILIPIGI NICAPGODELLALANNESLEDEVRDIRS NHHLOGKLEDPAIRWOMALVKDLPNETDDTS DPEKTYER VLDIANVLFHLEGKEVGEVERPERIY CLVEHPORSKKAVWHKLLSKORKRAVVACF RMAFLYNIPRIRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKFGAEPPEEDGTKRAVDEH ULLIFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDOEEEVKSFFEKEMEKOKLLVQQ ARLHDRGAAEMVLQTISASKGETGFMVAAT LKLGIALINGONSTVQOKMLDVLKEKKDVGF FOSLAGLIMQSCSVLDLNAFERONKAGGLGM VTEGGGEKVLQDDEFTCLFFLQLLCEGH NSDFQNYLRTOTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEGGORPISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOQRNISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOQRNISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOLLEKEMDLQ KDMVVWILLSMLEGNVVNGTIGKGMVDMLV ESSNNEMILKFFDMFLKLKDTSDTFKEYD PPOKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVRFHEPARDIGFRNA VILITLISEHMFNDTRLQTIFLAESVLNYFQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKSSKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLINGSANKEESEKERPEE QGFRAAFFSLLTVRSALFALRYILAMBALS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFGFFRIICSLLIGGSLVEGA KKKVAELLANMPDPTQDEVRGDEEGERKP LEAALPSSDLTDILKELTEESDLSDIFGLDLKR EGGQYXLIPHNPNAGLSDLMSNPVPMPEVQE KFOGKA KEEKEKEKETSISTESLSDIFGLDLKR EGGQYYLIPHNPNAGLSDLMSNPVPMPEVQE KFOGKA KEEKEKEKEETSEAGEDGE KEEKAKEDKGKOKLRQLHTHRYGEPEVPSA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGRELFRASSSENA KVYSLDSSSRRIAVHYVLESSGYMEPTVRL PILHTVISFCGTGYYCLKVPLVIKREKEKEVARK LEFOGLYTTEQFSEDDIKGQWDRI.VINTGSPF NNYWDKFVKRKVMDKYGEFRUFEITFFFVVI CHANNOFFFAAHLLDLAMGFKTLKTILSSVTH NGKQLVLTVGLLAVVVLLYFARNFERKF YNSSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGIGDEIEDPAGDEYSTVATLPHTYFFVV VILLAILQGLIIDAFGELRDQQEQVKEDMETKC FICMGINDYFDTVPHOFETHTLQERNLANVFL FILMYLINKDETEHTGGESTVWKMYYGERCWE FFRAGDCFRKQYEDQUN  AGELLINSDDPPALASOS,GRITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SEPRSWDYRRWPHPHANFCFCRDESLAMML  5552 AGELLINSDDPPALASOS,GRITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SEPRSWDYRRWPHPHANFCFCRDESLAAML SEPRSWDYRRWPHPHANFCFCRDESLAAML SEPRSWDYRRWPHPHANFCFCRDESLAAML	1 -			Ì			KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
ERKKMERGDRYSMOTSLIVAALKRILIPIGI NICAPGODELLALANNESLEDEVRDIRS NHHLOGKLEDPAIRWOMALVKDLPNETDDTS DPEKTYER VLDIANVLFHLEGKEVGEVERPERIY CLVEHPORSKKAVWHKLLSKORKRAVVACF RMAFLYNIPRIRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKFGAEPPEEDGTKRAVDEH ULLIFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDOEEEVKSFFEKEMEKOKLLVQQ ARLHDRGAAEMVLQTISASKGETGFMVAAT LKLGIALINGONSTVQOKMLDVLKEKKDVGF FOSLAGLIMQSCSVLDLNAFERONKAGGLGM VTEGGGEKVLQDDEFTCLFFLQLLCEGH NSDFQNYLRTOTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEGGORPISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOQRNISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOQRNISKALQVA KQVPNILTEYIQGPCTGNQQSLAHSRLWDAV VGFHVFAHMOMKLSQDSSOLLEKEMDLQ KDMVVWILLSMLEGNVVNGTIGKGMVDMLV ESSNNEMILKFFDMFLKLKDTSDTFKEYD PPOKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVRFHEPARDIGFRNA VILITLISEHMFNDTRLQTIFLAESVLNYFQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKSSKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLINGSANKEESEKERPEE QGFRAAFFSLLTVRSALFALRYILAMBALS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFGFFRIICSLLIGGSLVEGA KKKVAELLANMPDPTQDEVRGDEEGERKP LEAALPSSDLTDILKELTEESDLSDIFGLDLKR EGGQYXLIPHNPNAGLSDLMSNPVPMPEVQE KFOGKA KEEKEKEKETSISTESLSDIFGLDLKR EGGQYYLIPHNPNAGLSDLMSNPVPMPEVQE KFOGKA KEEKEKEKEETSEAGEDGE KEEKAKEDKGKOKLRQLHTHRYGEPEVPSA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGRELFRASSSENA KVYSLDSSSRRIAVHYVLESSGYMEPTVRL PILHTVISFCGTGYYCLKVPLVIKREKEKEVARK LEFOGLYTTEQFSEDDIKGQWDRI.VINTGSPF NNYWDKFVKRKVMDKYGEFRUFEITFFFVVI CHANNOFFFAAHLLDLAMGFKTLKTILSSVTH NGKQLVLTVGLLAVVVLLYFARNFERKF YNSSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGIGDEIEDPAGDEYSTVATLPHTYFFVV VILLAILQGLIIDAFGELRDQQEQVKEDMETKC FICMGINDYFDTVPHOFETHTLQERNLANVFL FILMYLINKDETEHTGGESTVWKMYYGERCWE FFRAGDCFRKQYEDQUN  AGELLINSDDPPALASOS,GRITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SEPRSWDYRRWPHPHANFCFCRDESLAMML  5552 AGELLINSDDPPALASOS,GRITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SEPRSWDYRRWPHPHANFCFCRDESLAAML SEPRSWDYRRWPHPHANFCFCRDESLAAML SEPRSWDYRRWPHPHANFCFCRDESLAAML	I				1		
NICAPGDOELLALAKNITSLKDTEDEVPIDIRS NIHHQGKLEDPARWQMALYKDLPNKTDDTS DPEKTVERVIDLANVLFHLEQKSKRVGRRHY CLVEHPQRSKKAVWHKLLSKQRKRAVVACF RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKRGAEPPEEDEGTKRVDHO LILLFSRTALTEKCKLEEDPLYMAYADIMAKS CHDEEDDDOEEEVKSFERKEMEKQKLLYQQ ARLHDRGAAGMUQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDVLEKKRVOGF PQSLAGILMQSCVULDNAFERQNKAEGLGM VTEGSGEKVLQDEPTCDLFEFLQLLCEGH NSDFONVLRIOTGONTTVNIHTVYULRVQ ESISDFYWYYSGKDVUDEOGQRNFSKAIQVA KQVFNTLTEYIQPETCHOOLAGNEW KQVFNTLTEYIQPETCHOOLAGNEW SOMVVMLISMLEGNVVNGTIGKQWYDMLV ESSNNVEMILKFDMFLKLANBLWDAV VGFLHVFAHMQMKLSQDSSQIELKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQWYDMLV ESSNNVEMILKFDMFLKLANBLWDAV VGFLHVFAHMQMKLSQDSSQIELKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQWYDMLV ESSNNVEMILKFDMFLKLANBLWDAV VLLTNLSEHMPNDTRLQTFLELAESVLNYPQP PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNVA VLLTINLSEHMPNDTRLQTFLELAESVLNYPQP FIGGREMGSAKRIERVYFEISESSRTQWEKPQ VKESKROFIEDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNEFSSRNTWEED GPRMAFFSILTVRSALFALRYNTTLMRMLS LKSLKKQMKKVKKMTVKAMTAFFSSVWSI PMTLHFVASVPRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDEEGERKP LEAALPSEDLIDLKELTEESDLSDIFGIDLIKR EGGQYXLIPHNPNAGLSDLMSNFPVMFEVQE KFQCKAKEEFKEEKEETISELSDLSDIFGIDLIKR EGGQYKLIPHNPNAGLSDLMSNFPVMFEVQE KFQCKAKEEFKEEKEETISELSDLSDIFGIDLIKR EGGQYKLIPHNPNAGLSDLMSNFPVMFEVQE KFQCKAKEEFKEEKEETISETSSGVMPEYVPSA AFAINFILLFYKVSTSSVVEGKELPTRISSENA KVTSLDSSNFRIAVHYVLEESSGYMPEPVRIL PLHTVISFFCIIGYYCLKVPLVFRKREKEVARK LEFDGLYTTEOPSEDDIKGQWDRI LINTOSFP NNYWDKFVKRKVMDKYGEFYGRDRISBLIA MDKAALDFSDAREKKFRKDSSLSAVLNSID VKYQMWKLGVVFTNDNSFLYATGHTFFLY VILLAIQGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFFV VILLAIQGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFFV VILLAIQGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFFV VILLAIQGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFFT VILLAGGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFFT VILLAGGGLIDAFGELRDQQEQVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFTV VILLAIQGLIDAFGELRDQQEGVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFTV VILLAIQGLIDAFGELRDQQEGVKEDMETKC FICGGRONYPTVPHOFFETHDTTFFTV VILLAIQGLIDAFGELRDQQEGVKEDMET	Į.	l				ł	` ` `
NHILOGKLEDPAIRWOMALYKDLPNRTDDTS DPEKTYER VIDLANVI-FILEGKRYGERHY CLVEHPORSKKAVWHKLLSKQRKRAVYACF RMAPLYNIPRIRAVNI-ELQGYERSWIETEEH YFEDKLIEDLAKFGAEPPEEDGTKRVDPLHO LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDCEEVKSFERERKGKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIALINGGNSTVQCKMLDVLKEKKDVGF FQSLAGILMQSCSVLDLMAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCLFFLOLLCEGH NSDFQNYLRTQTGNNTTYNIIISTYDYLLRVQ ESISDFYWYYSGKDVIDEGGGRNFSKARQVA KQYFNITE-BYQPOFTGNQSLAHSRLWDAV VGFLHYFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTKEMDL KDMVVMLLSMLEGNVVNGTKEMDL ESSNNVEMILKFDMFLKLKDLTSSDTFKEYD PDGKGVFKRPFHRAMESHKHYTQSETEFLL SCAETDENETLDYEEFVRFHEPAKDIGRNVA VILITALSHMPMOTRICTHE LEVYLNYYOF FLGRIEMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVNNEGGREKMELFVNNCED TIFFBMQLAAQISSDLNERSANCESEEKERPEE QGRMAFFSILTVRSALFALRYNNITLMRMLS LKSLKKQMKKVKMTVKDMVATAFFSSYWSI FMTLLHFVASVFRGFFRIGSLLLGGSLVEGA KKKVAELLAMMPDPTQDEVRGDEEGEREKERE QGRMAFFSILTVRSALFALRYNNITLMRMLS LKSLKKQMKKVKMTVKDMVTAFFSSYWSI FMTLHFVASVFRGFFRIGSLLLGGSLVEGA KKKVAELLAMMPDPTQDEVRGDEEGEREKERE EGGGYKLIPHPNAGLSDLMSRNPVPMPEVQE KFQQKAKEEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPSA FWKKIIAYQQKLINTYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELFKRSSENA KVYSLDSSKRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGFYCLKVFLUFKREKEVARK LEFGGLYTIEPSEDDIKGGWALIVITGSSPF NNYWDKFVKRKVMWKYGEFYGRDRISELLG MDKAALDFERDAERKKEPKEVARK LEFGGLYTLEPSDEDIKGGWALIVITGSSPF NNYWDKFVKRKVMWKYGEFYGRORISELLG MDKAALDFERDAERKEPKEVGSSLSAVLNSSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHTYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLJTVVAFNFFRFK YNKSEDGDTPDMKCDDMLTCYMFHWYVGV RAGGGGDELEDRAGDEVSTYRIEDTITFFFV VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICIGNDYPTDVPHOFFETHTYVGV RAGGGGDEDEDRAGPCSPLAMML FILMYLNKDETEHTGQCSYVWKMYQERCWE FFPAGDCFRKQYEDQLN.  446 1796 A 3592 1 355 AGLELINSDPPALASQSAGITGVTRTPSLFF+ DTVLLCCSGWSAVAPSRLTAALFS*NQAVCL SEPRSWDYRRWPHPHANFCFCDDEJSLAMU. SEPRSWDYRRWPHPHANFCFCDDEJSLAMU. SEPRSWDYRRWPHPHANFCFCDEJSLAML	ł			ľ		ļ	,
DPEKTVERVLDIANULFHLEGKSKRVORRHY CLVEHPORSKKAVWHKLISKORKAVVACE RMAPLYNLPRHRAVALFL GGYEKSWETEEH YFEDKLIEDLAKPGAEPPEDEGTKRVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDOGEEVKSFERKEMEKOKLLYQQ ARLHDRGAAGMVLQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDVLKEKKDVGF FOSLAGIMOSCSVLDLNAFERONKAEGLGM VTEGGGEKVLQDDEFTCDLFRFLQLLCEGH NSDFONVLRIOTGONTTVNITDVLLRVQ ESISDFYWYYSGKDVUDEGGGRNSKARQVA KQVPNTLESTYGOPCTONGOSLAHSRLWDAV VGFLHYFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFDMFLKLANDLSSDTFKEYD PDGKGVIFKRDFHKAMESHCHYTQSETIEFLL SCAETDENETLDVEEFVKRFEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESSVLNYPQP FLGRIEMGSAKRENYYFEISSTYQMEKPQ VKESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSARGEVFYFEISSTYQMEKPQ VKESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSARIQWAK VULTHLSEHMPNDTRLQTFLELAESSVLNYPQP FLGRIEMGSAKRENYYFEISSTYQMEKPQ VKESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSARIQWAK VLLTHLSEHMPNDTRLQTFLELAESSVLNYPQP FLGRIEMGSAKRENYYFEISSTYQMEKPQ VKESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSARIQWEK VKESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSARIQWEK KESKRGFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKESKRQFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKESKRQFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKESKRQFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKESKRQFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKESKRQFIFDVVNEGGEKEKMELFVNYCED TIFEMQLAAQISSSDLNERSKYQMERD VKSCRQFIFDVVNEGGEKEKMETSSEKGEFERE GGGYKLHPNNPNAGLSDLNSNYPVMERPEVESA FWKKILHYVKSTSSVVEGKELFTRSSEKAREGDGG KEEKAKEDKGKULNYTARNFYNMRMLALFY AFAINFILLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSNIRIAMYTYLEESSGYMEPTVRIL VTLJHTVISPFCIIGYYCLKVPLVYFKREKEVARK LEFDGLYTTENGPEDDIKGQVDLVINTOSSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAAEKKFRKOSSLSAVLNSID VKYQMWLLGVVVTYNNSFLUTINTOSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAAEKKFRKOSSLSAVLNSID VKYQMWLLGVVVTTNNSFLYTYVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYFUGV RAGGIGDEIDEPAGDEVSTYNTYNTYNTH NGKQLULTVYQLENLANVLT FLMYLINKOETEHTGQESYVWKMYQERCWE FFPAGDCFRCYFEDDIKGOMLTCY FFPAGDCFRCYFEDDIKA	1					ļ	
CLVEHPORSKKAVWHKLLSKÖRKRAVVACE RMAPLYNLPRHAVNLFLÖGYKESWETEEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEVKSFEEKMEKKÖKLLYQQ ARLHDRGAABMVLQTISASKGETGMVAAT LKLGIAILNGONSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERONKAGGGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCGGH NSDFQNVLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYSCKDVIDEQONPSKALQVA KQYFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQELLKELMDLQ KDMYVMLLSMLEGNVVNOTIGKQMVDMLV ESSNNVEMILKFEDMFLKALDLTSSDTFKEYD PDCKGVFKRPDHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRPHEPAKDIGFNVA VLLTNLSEMHMYDITRLGTELAESVLNYPQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDGLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDGLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDGLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDGLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDGLVEGA KKIKVAELLANMPDFTDGEVEGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIGLDLKR EGGGYKLIPNFNAGLSDLMNPYPMPEVEV EGGGYKLIPNFNAGLSDLMNPYPMPEVE EGGGYKLIPNFNAGLSDLMNPYPMTEV AFANFILLFYKVSTSSVVEGKELFFTKSSENA KVTSLDSSSHRIAVHYVLEESSGYMEPTIVRIL PILHTVISFFCIIGYYCLKVPLVIVIKRSSENA KVTSLDSSSHRIAVHYVLEESSGYMEPTIVRIL PILHTVISFFCIIGYYCLKVPLVIVIKRSSENA KVTSLDSSSHRIAVHYVLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCTMFHMYVGV RAGGGIGDEEDPAGDEVTYRILDSVTH NGKQLVLTVGLLANVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCTMFHYVGVE VILLAIIQGLIDAFGELRDQQEVKEDMETKC FICGIGNDYFDTVPHGETHTILOEHLANYLC FLOGIGIDAFGELRDQQEVKEDMETKC FICGIGIGDLFPAGESTYWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*QAVCL SLPRSWDYTRWPPHPANFGCFCRDESLAML		]				l	1 1
RMAPLYNLPHRAVNLFLQGYEKSWETEEH YFEDKLIEDLAKPGAEPPEEDEGTKR VOPDHO LILLFSRTALTEKCKL EEDFLYMAY ADIMAKS CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAABMVLQTISASKGETGPMVAAT LKLGIAILNGGNSTVQQKMLDYLKKKDVGF FQSLAGLINGSCSVLDLNAFERQNKAGGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNVLRTQTGNNTTNISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKALQVA KQVFNTLTEYIQGPCTONQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGRVVNOTIGKQNVDMLV ESSSNVEMILKFFDMFLKLKDLTSSDTFKEYD PDCKGVFKRDFHKAMESHKHVTQSTFEFL SCAETDENETLDYEEFYKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYPQP FLGRIEIMGSAKRIERVYFEISESRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESKERPEE QGFRMAFFSIT LVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICALLGGSLVSGA KKIKVAELLANPPDPQDEVRGDGEGERKP LEAALPSEDLTDLKEETESDLLSDIFOLDLKR EGGGYVKLIPHNENAGLSDLMSNFYPMEEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKOKQKLRQLHTHRYGEPEVYESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFABFELLFYKVSTSSVVEGLPTRSSSENA KVTSLDSSHRIIAVHYVLEESSGYMEPPVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYTIEPGSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRGRISELLG MDKAALDFSDAREKKRYKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFRAHLDLINGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRGRISELLG MDKAALDFSDAREKKRYKLDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFRAHLDLINGQPDRLVNTTQSFP NNYWDKFVKRCKVDDLTCVMFTIMYVGU RAGGGIGDEEDPAGDEFYRILBTILTSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCVMFTIMYVGFP VILLAIIQGLIDAFGELRDQQQVERDMTKC FICGIGNDYPTOTVHGFFTHILQEENLANNYLF FLMYLINKDETEHTIGGESYVWKMYQERCWE FFRAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALLFS*AQAVCL. SLPRSWDYTRWPPTPRAFFICFCRDSSLAML	1			'		ļ	
YFEDKLIEDLAKPGAEPPEDEGTKVDPLHQ LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGMVAAT LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKUQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTOTONTTVMIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGRNFSKAIQVA KQVFNILTEYIQGFCTGNQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDML ESSNNVEMILKFEDMFLKLKDLTSSDTFKEYD PDGKGVFKRDPHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VILITNLSEHMPNDTRLQTFLELAESTLNYFQP FLGREIMGSAKRIERVYPESESSSTOWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERANKEESKERPEE QGRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHIFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDTQDEVRGDGEGERKP EGGQYKLJPHNPNAGLSDLMSNPYMPEVQE KFQEKAKEDKGKKQKLRQLHTHRYGEPEVPESA FWKSILAYQQKLLNYFANFYNMRMLALFV AFANFILLFYKVSTSSVEGKELFTRSSSENA KVTSLDSSSHRIJAVHYVLEESSGMEFTVRIL PILHTVISFFCIGYYCLKYPLJVFKREKEVARK LEFDGI YITEOPSSDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMKYGEFYGRDRISELLG MDKAALDFSDAREKKFKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFRAHLLDLINGFULTSVYTH NGKQLVLTVGLIAAVVVLYTVVAFNFFRKF YNKSEDGDTFDMKCDDMLTCTMFHMYVGV RAGGGIGDEIDPAGGESTVUKMYGERCWE FFFAGDCFRAHULDLAVVVLYTVVAFNFFRKF YNKSEDGDTFDMKCDDMLTCTMFHMYVGF RAGGGIGDEIDPAGGESTVWKMYGERCWE FFFAGDCFRAKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFSPAQAVCL. SLPRSWDYYRWPPTPRAFFCFCRDSSLAML						1	
LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAABMVLQTISASKGETGPMVAAT LKIGIAILNGGNSTVQKMLDYLKEKDVGF FOSLAGLMOGSCSVLDLNAFERONKAGGIGM VTEEGSGEKVLQDDEFTCDLFFFLQLCEGH NSDFQNYLRTOTGNNTYNISISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGGRNFSKAIQVA KQVFNTLTEYIQGPCTGNQGSLAHSRLWDAV VGFLHYFAHMQMKLSQDSSQIELIKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVFKRDFHKAMESHKHYTGSETFFLL SCAETDENETLDYEFFYKRFHEPAKDIGFNVA VLLTNI.SEHMPNDTRLQTFLELAESVLNYPQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKESRROFHFDVVNEGGKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESSKERPEE QGFRMAFFSIT LVRSALFAHRYMILTLMRMLS LKSLKKQMKKVKKMTVKDMYTAFFSSYWSI FMTLLHFVASVFRGFFRLGSLLAGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGDLLKR EGGQYKLIPHNPNAGLSDLMSNPYPMEEVQE KFQEQKAKEEKEEKEETKSEPKAGEDGE KEEKAKEDKOKGKLRQHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFARPELLFYKVSTSSVVEGELPTRSSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKKEEVARK LEFDGLYTTEGPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRENISELLG MDKAALDFSDAREKKRYKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMIMSVL GHYNNFFFRAHLDLINGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRENISELLG MDKAALDFSDAREKKRYKLDSSUSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMIMSVL GHYNNFFFRAHLDLINGGWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRENISELLG MDKAALDFSDAREKKRYKLTRILSSVTH NGKOLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPJDMKCDDMLTCYMFFHNYVGK PKRSGGGIGBEEDPAGDEFYRIETHITJCEFHNYVGY RAGGGIGBEEDPAGDEFYRIETHITJCEFHNYVGY RAGGGIGBEEDPAGDEFYRIETHITJCEFHNYTKF FLMYLINKDETEHTGGESYVWKMYQERCWE FFFAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL. SLPRSWDYSWAWAPSRLTAALFSPAQAVCL.	1	•			·	İ	· · ·
CHDEEDDDGEEEVKSFEEKEMEKOKLLYOQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKIGIALINGONSTVQKMLDYIKEKKDVGF FQSLAGLMQSCSVLDLNAFERONKAEGLGM VTEEGSGEKVLQDDEFTCDLFRIQLICEGH NSDFQNYLRTOTGNTTVNIIISTVDYLLRVQ ESISDFYWYSGKDVIDEQGQRNFSKAGVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELIKELMDLQ KDMVVMULLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVFKRDFIKKANFLYSSETFFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLITNLSEHMPNDTRLQTFLELAESVLNYFQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNII.TLMRMLS LKSLKKOMKKVKKINTAMVATAFFSSWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVSRGEEGEKRKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGL.SDLMSNPYPMPEVQE KFQFQKAKEEKEKEETKSSPEKARGEDGE KEEKAKEDKGKCKLRQLHTHRYGEPEVPPSA PWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIJAVHYVLEESSGYMEPTVRIL PLHTVISFFCIGTYCLEVJUFTREKEEVARK LEFGGLYITEQPSEDDIKGQWDRLVNTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAAEKKRKFKKDSSLSAVLNSID VKYQMWKLGVVFTDNSSLVLAWYMMSVL GHYNNFFRAHLLDLAMGFKTLRTLSSVTH NGKQLVLTVGLLAVVVYLVTVAFNFFRKF YMKSEDGDTPDMKCDDMLTCYMFFHMYVGV RAGGGIGDEIEDPAGDEYIYRIFDITFFFVI VILLAIQGLIIDAFGELRQQQGVEDMETKC FICGIGNDYFTOTPHTQETTHLQHINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFFAGDGFRKQYEDOLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF+ DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPAGFICRCDESSLAML		į					· · ·
ARLHDRGAAEMVLQTISASKGETÖPMVAÄT LKLGIAILNGGNSTVQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLAAFBRQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRIQLLCEGH NSDFQNYLRTOTGNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNILTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTYKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNI.SEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKROFIFDVNNEGGGKEKKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFROFFRIICSLLIGGSLVEGA KKIKVABLLANMPDTQDEVRGDGEGEGRKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEEKETKSEPFKARGEDGE KEEKAKEDKGKQKLRQLHTHRYGEFEVPESSA FWKILIAYQQKLLNYFARNFYNMRMLALFV AFANIFILLFYKVSTSSVVEGKELPTRSSSENA KVSLDSSSHRIJAVHYVLEESSGYMEPTVRL PLHTVISFFCIIGYYCLKVPLVIFREEKEVARK LEFGGLYTTEQFSEDDIKGQWNDTLVINTQSFP NNYWDKYVRKVWNDKYGEFFYGRDRISEBLLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNIFFRAHLLDLAMGFKILRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGOTPDMKCDDMLTCYMFTHMYVGV RAGGGIGDEIEPAGDEYEYRIEIDITTFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDVFDTVPHIGFETHTLQEINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFFRGGOFRRQYEEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF+ DTVLLCCSGWSAVAPSLTAALFS*AQACU- SLPRSWDYRRWPPRAPAFCIFCRDESSLAML	]				]	Į.	LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDPQNYLRTQTGNNTTVNILISTVDYLLRVQ ESISDFYWYSGKDVIDEQGQRNFSKALQVA KQVPNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFILIVFAHMQMKLSOSSOIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFIKKABKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLITILSEHMPNDTRLOFLELAESVLNYPQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRGFIFDVVNEGGEKKKMEKFVNFCED TIFEMQLAAQISESDLNSRSANKEESKERPEE QGPRMAFFSILTVRSALFALRVNILITLMRMLS LKSLKKQMKKVKKMTVAMVTAFFSSYWSI FMTILHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKFEKKEEKETKSSPFKARGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIJAVHYYLEESSGYMEPTVRLI PILHTVISFFCIGTYCLEVJFFREKEVARK LEFGGLYTTEQPSEDDIKGQWDRLVNTQSFP NNYWDKFVRKVMVNGYEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLVLAWYMTMSVL GHYNNNFFRAHLLDLAMGFKTLRTLSSVTH NGKQLVLTVGLLAVVVYLVVVAFNFFRKF YNKSEDGDTPDMKCDULTCYMFHRWYVGV RAGGGIGDEEDPAGDEYEIYRIFDITFFFVI VILLAIQGLIIDAFGELRDQQEQVEEDMETKC FICHGIGNDYFDTVPHIGFTETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFRAGDCFRQYEEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPAGFICRCDESSLAML		].					CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ
LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDPQNYLRTQTGNNTTVNILISTVDYLLRVQ ESISDFYWYSGKDVIDEQGQRNFSKALQVA KQVPNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFILIVFAHMQMKLSOSSOIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFIKKABKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLITILSEHMPNDTRLOFLELAESVLNYPQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRGFIFDVVNEGGEKKKMEKFVNFCED TIFEMQLAAQISESDLNSRSANKEESKERPEE QGPRMAFFSILTVRSALFALRVNILITLMRMLS LKSLKKQMKKVKKMTVAMVTAFFSSYWSI FMTILHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKFEKKEEKETKSSPFKARGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIJAVHYYLEESSGYMEPTVRLI PILHTVISFFCIGTYCLEVJFFREKEVARK LEFGGLYTTEQPSEDDIKGQWDRLVNTQSFP NNYWDKFVRKVMVNGYEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLVLAWYMTMSVL GHYNNNFFRAHLLDLAMGFKTLRTLSSVTH NGKQLVLTVGLLAVVVYLVVVAFNFFRKF YNKSEDGDTPDMKCDULTCYMFHRWYVGV RAGGGIGDEEDPAGDEYEIYRIFDITFFFVI VILLAIQGLIIDAFGELRDQQEQVEEDMETKC FICHGIGNDYFDTVPHIGFTETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFRAGDCFRQYEEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPAGFICRCDESSLAML	1	1	1		!	(	ARLHDRGAAEMVLQTISASKGETGPMVAAT
FOSLAGLMSCSVLDLNAFERONKAEGIGM VTEGSGEKVLODDEPTCDLFREJOLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNNTLEFYIQGFOTGNQGSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLIJVEFEVKRFHEPAKDIGFNVA VLLTNLSBHMPNDTRLQITLELAESVLNYFQP FIGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIEDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTVSALFALRYNILTILMRMLS LKSLKKQMKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDFTQDEVRGDGEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQVXLIPHNPNAGISLDLMSNPYPMPEVQE KFQEQKAKEEKEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPPEVPESA FWKKILAYQKLLNYFARNFYNNRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGGWDRJ.VINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFNDSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMVGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFIMYVGF RAGGGIGDEIEDPAGDEVEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FIGGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFAGGOCRK, OKAVCL SLPRSWDYRKWPHPANICIFCRDESLAML	1	1				ł	1
VÎTEGGGEKVLQDDEFTCDLFRILQLLCEGH NSDFQNYLRTQTONTITVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGFCTGNQOSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDAU, KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVJFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKFHEPARDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEMELFYNFCED TIFEMQLAAQISESDLNERSANKEESSKEPPEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKVKKNMTVKDMVTAFFSSYWSI FMTLLHFVASVFGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGGERKP LEAALPSEDLTDLKETESDLLSDIFGLDLKR EGQQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKETKSFEKAEGEDGE KEEKAKEDKGKCKLQHTHRYGPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVFKREKEVARK LEFDGLYTEOPSEDDIKGQWDRLVNTNOSPP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKFKKKDSSLSAVLNSID VKYQMWKLGVVTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGF RAGGIGDEIEDPAGDEVSTRIFDITFFFFVI VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYQERCWB FFFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYQERCWB FFFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYQERCWB FFFAGGCFRKQVEDQLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYQERCWB FFFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWB FFFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGELRDQOEQVEEDMETKC FFAGGCFRKQVEDDLIN VILLAIIQGLIDAFGPLRDQOEQVEEDMETKC FFAGGCFRKQVEDDLIN VILLCCSGWSAVAPSRLTAALPS*AQAVCL SLPRSWDYRRWPHPANFCIFCRDESLAML	1	i					
NSDFQNYLRTQTGNTTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQYFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMYVMLLSMLGNVNGTIGKGWDMLV ESSNIVEMILKFDMFLKLKDLTSSDTFKEYD PDGKGVIFKDFHKAMESHKHYTQSETEFLL SCAETDENETLDYBEFVKRFHEPAKDIGFNVA VILTINLSEHMPNDTRLQITLELAESVLNYFQP FLGREIMGSAKREVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKESSEKERPEE QGPRMAFFSILTVSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI PMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLKELTEESDLSDIFGLDLKR EGGQVXLIPHNFNAGLSDLMSNFVPMPEVQE KFQEQKAKEEEKEEKESTKSFEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMSMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTVRIL PLHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLVITTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFUNDSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGE RAGGIGBEIEDPAGDEVEIYRIFDITFFFFVI VILLAILQGLIIDAFGELRDQQEQVKEDMETKC FIGGIGNDYFDTVPHGFETHTLQEHNLANYLI FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQVEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF+ DTVLLCCSGWSAVAPSRLTAALPS*AQVCL SLPRSWDYRKWPHPANFCIFCRDESLAMUL SLPRSWDYRKWPHPANFCIFCRDESLAML	1	<b>!</b>					, , ,
ESISDFYWYYSGROVIDEGORNFSKAIQVA KQVFNTLTEYIQGPCTONQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFPDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNVA VLLITNLSEHMFDTLQTIFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKMTVKDMVTAFFSSYBS FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDFTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYTTEQPSEDDIKGQWDRILVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVYTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLIAVVVYLYTVVAFNFFRKF YNKSEDGDTFDMKCDDMLTCYMFHMYVGV RAGGGIGDEEDPAGDEYFIYRIIEDTITFFFFY VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTIGQESYVWKMYQERCWE FFPAGDCFFRKYGEDQLN  446 1796 A 3592 1 355 AGELLNSDDPPALASQSAGITGVTRTPSLFF+ DTVLLCCSGWSAVAPSITAALFS*AQAVCL SLPRSWDYRRWPHPANFCFCRDESLAML	1	í	Ì	ŀ		1	
KQVFNILTEYIQOPCTONQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTTKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKESSEKEPFEE QGPRMAFFSLTVRSALFALRYNLTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHFVASVPRGFFRICSLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGGERFY LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAREDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYLLKYPLIVFREKEEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVRRKVMDKYGEFYGRRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLDLIAMGFKTLKTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYJRIIEDTIFFFFY NKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEEDPAGDEYJRIIEDTIFFFFYFFFFTPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDPPALASQSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPHPANFCIFCRDESLAML	1	ļ					1 ' '
VOELHYFAHMOMKLSODSOGIELKELMDLQ KDMVVMLLSMLEGNVVNGTIGKOMVDMLV ESSNNVEMILKFPDMFLKLKDLTSSDTFKEYD PDGKGVFRRDFHKAMESHKHYTOSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPDDTRLQTFLELAESVLNYFQP FLGRIEINGSAKRIERVYFEISESSRTQWEKPQ VKESKROFIFDVVNEGGEKEKMELFVNFCED TIFFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKOMKKVKKMTVKDMVTAFFSSYWSI FMTLHFVASVFRGFFRIICSLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLINYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVAKK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKEPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVARNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGIGDEIEDPAGDEYEIYRIIFDITFFFFFFF VNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGIGDEIEDPAGDEYEIYRIIFDITFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	}	İ					
KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRPHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSKTQWEKPQ VKESKROFIEDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKEPPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKVKKMTVKDMVTAFFSSYWSI FMTLHFVASPEGFRICISLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYKLIPHNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFANRYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIGYYCLKVPLVIFKREKVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDISEELIG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL, GHYNNFFFAAHLLDLAMGFKITRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEGDDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYPDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGTTGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLFPSWDYRRWPPHPANFCIFCRDESLAML	1	]		· '			7
ESSNNVEMILKFFDMELKLKDLTSSDTFKEYD PDKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPMDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKROFIEDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKKVAELLANNMPDTQDEVRGDEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSESMA KVTSLDSSSHRIAVHYVLEESGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLVITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRVMDKYGEFYGRDRISSELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFITDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKILRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYETYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFTHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGEELLNSDDPPALASQSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLFRSWDYRRWPPHPANTCIFCRDESLAML	}	1	}	1			
PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VILITNLSEHMPNDTRLQITLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFIEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVSVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFGGLYITTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKFVKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEEDPAGDEVEIVRIIFDITFFFFI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	J	}	l			1	
PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VILITNLSEHMPNDTRLQITLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFIEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVSVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFGGLYITTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKFVKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEEDPAGDEVEIVRIIFDITFFFFI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	1					<b> </b>	ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD
SCAETDENETLDYBEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTIFLEAESVLNYFQP FIGRIEIMGSAKRIERVYPEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTIMEMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQECKAKEEKEEKEEKTSEPEKAEGEDGE KEEKAKEDKOKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKQQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEEDPAGDEVIRVIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLMYLINKDETEHTGGESSYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	i						
VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGREIMGSAKRIENYYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFYNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHIFVASVYFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKURQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGTIGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESISAML	1	1				1	1
FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQVEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	}	1	1	ĺ			1
VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE QQFRMAFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETIKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIIFDITFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDESLA/ML	ſ	[				[	1
TIFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTVRSALFALRYNILTIMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIIGSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEFTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTIRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDESLA/ML	1		ľ	ľ			, , ,
QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGGYKLIPHNPNAGLSDLMSNPVPMPEVQE KFOEQKAKEEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFKFF YNKSEDGDITPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLA/ML	1	}	i	i		<b>!</b> .	1
LKSLKKQMKKVKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML	1			!		1	1
FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERRP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEEKESFEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML	•	1	[	<b>1</b> .		[	
KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	}	}		1	ł	
LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKCSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1		1		ł	
EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQHTHRYGEPEVPESA FWKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1		l			KKIKVAELLANMPDPTQDEVRGDGEEGERKP
KFQEQKAKEEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMI.ALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLK VPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1	1	(	1	l	LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
KFQEQKAKEEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMI.ALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLK VPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1.	1	1	1	}	1	EGGQYKLIPHNPNAGLSDLMSNPVPMPEVOE
KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTQEHNLANYLF FLWYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1			}		
FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1	1	1			
AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1	l	1		1	
KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	J	ļ	j	ļ	1	
PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML			ł	j		]	
LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYVNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	Í	1	· '	1			
NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1	]			1	
MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	!	<b>]</b>	J	}	J	
VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHY\NNFFRAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	i		1		1	
GHYNNFFRAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		l	1	l	•	İ	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
GHYNNFFRAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1	1	1	l	1	VKYQMWKLGVVFTDNSFLYLAWYMTMSVL
NGKQLVLTVGLLAVVYYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	]		1		1	1	1 -
YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1	i	1	1	1	
RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	l	1	Ì	ł	1	
VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1	}	J	}	}	
FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	}	ĺ	]	}			· · · · · · · · · · · · · · · · · · ·
FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN  446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1		1		1		
FFPAGDCFRKQYEDQLN   446   1796   A   3592   1   355   AGLELLNSDDPPALASQSAGITGVTRTPSLFF*   DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL   SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	)	ļ	}	}	1	
446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	]	} .	} ·			
DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	İ		l	1		FFPAGDCFRKQYEDQLN
DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1		1	J	}	}	
			۱.	1			
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	İ	[	1	{	İ	
	L	ــــــــــــــــــــــــــــــــــــــ	└──	L	L	J	1

NO. of   No. of   N	250 15	CEO ID	3.6-4	CEO	Drediated	Deadisted and	Amino cold converse (AAloria G. C.
mucheotide cocide concide corresponding overseponding of peptide corresponding on the stamino acid residue of peptide sequence peptide sequence peptide sequence company to the stamino acid residue of peptide sequence peptide sequence peptide sequence company to the stamino acid sequence company to the stamino acid sequence company to the sequence peptide sequence peptide sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the sequence company to the sequence peptide sequence company to the	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Seq			noa				
sequence         uence         09/496 or personal in last amino acid residue of peptide sequence peptide sequence peptide sequence         m/methodine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, S=Serine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, C=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, C=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, C=Glitamine, R=Aspirangie, R=Protine, O=Glitamine, C=G			}				
uence   914   ng to first anino said residue of peptide residue of peptide sequence   Pep	1	•					l=Isoleucine, K=Lysine, L=Leucine,
mino acid residue of peptide residue of peptide sequence	seq-	uence					
residue of peptide sequence peptide sequence peptide sequence peptide sequence per	uence			914			
Pepside   Pepside   Pepsible nucleotide deletion, Pepsible   Pepsible nucleotide insertion   Pepsible   Peps					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1797   A   3598   1202   1070					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1797   A   3598   1202   1070	J				peptide	· ·	
1797   A   3598   1202   1070							
GR*V*GAAASQQA/GSLPRRTTGGBIFGAQ*   QVWGAAAVCAPAM*ANARAGBRAG*    NSPWRSGVLAARAVGAGPWP*PPGCA*    NSPWRSGVLAARAVGAGPWP*PPGCA*    NSPWRSGVLAARAVGAGPWP*PPGCA*    NSPWRSGVLAARAVGAGPWP*PPGCA*    SPSRAPGLASGPAPLLQGVISSAGPLC*    NGTLALGLRP**AWGWGWRPKG*    GVEPLPYNNYKEVEKHGPGKWVLAAVL    LLLVLLGIGILVWHLQYRDVRQKVYDAAVL    LLVSGVPFLGPYHKESA,VTAFSEGSVIAYV*    LLVSGVPFLGPYHKESA,VTAFSEGSVIAYV*    FSIPGHLVEEARVMAEREVYMLPPRARSI   FVVTSVVAFPTDSKTVQRTQDNSCSFGLH,   GVELMRPTTFGFPSPYPAHACKQWALRG   ABSVLSLTFRSFDLASCDERGRHLVTVYN   SPMEPHALVQLGGTYPPSYLTHISSQN   LITLITITITERRHPGGPEATFFQLPRMSSCGG   RAQGTENSPYPOHYPPHOTHISSQN   GRLARYLSYDSSDPCGGFTCRTGCREREI   CDGWADCTDHSDELNCSCDAGHQTECK   CEPLEWCDSLUDGGDSNSDEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSDSSEGGSCPAQ*    RCSNGKCLSKSQQCNGKDDCDGSSSEGGSSSEGGSSEGGSSSSSEGSSEGGSSSSEGGSSSSEGGSSSSEGGSSSSEGGSSSSEGGSSSSEGGSSSSEGGSSSSSS	447	1707	Δ	3508		1070	<b></b>
1798   A   3604   3115   557   FRRKGGGGPKDFQAGLKYNSRHEKYNGLI GVEFLPYNNVKEVEKHGFGRWYLLAGLKYNSTHEKYNGLI GVEFLPYNNVKEVEKHGFGRWYLLAGL LLLVLLIGIGILVHLQYRDVYQKVPNOY RITTNENFVDAYENSNSTEFVSLASKVKQDL LVSGYFLGYHKESAVTAFSGSVIAYYW FSIPQHLVEEAUTAFSGSVIAYYW FSIPQHLVEEAUTAFSGSVIAYYW FSIPQHLVEEAUTAFSGSVIAYYW FSIPQHLVEEAUTAFSGSVIAYYW FSIPQHLVEEAUTAFSGSVIAYYW FSIPQHLVEEAUTAFFQLPRMSSCGG RAQGTPSSYYPGHYPPNIDCTWNEVPN QHYKVRFRFYLLEPGYPAGATCPKDYVEED EYCGERSGYVTSNNKHTVRFHSDQSYT GFLASYLSYDSSDPCRGGTCTTGTCRIER CDGWADCTDHSDELNCCGDAGDGTCKN CKPLFWVCDSLNDCGDNSDEQGCCPTATGCRIER CDGWADCTDHSDELNCCGDAGDGTCKN CKPLFWVCDSLNDCGDNSDEGGCSPASK KVNVVTCTKHTYRCLNGLCLSKGNPECD ECSDSDEKDCDCGLRSTFAGATCH CDGWADCTDHSDELNCCGDAGNEVGGT ADBGEWPWQVSLHALGGHICGASLISWN VSAAHCYDDDGFRYSDPTOHAFIGLAID QRSAPGVQERILKBLISIPFFNDFTFDVDIA ELEKPAEYSSMVRPICLPDASHYPFAGAT TOWCHTYGGTGALLCKGGREVNNGTT TULPQQTTPRMMCVGFLSGGVDSCQGDSG SVAAHCYDDGAGVSWGDGAGSAGS KVNVVTCTKHTYRCLNGLCLSKGNPECD GESCDSGDEKAGNAVGGT ADBGEWPWQVSLHALGGGHICGASLISWN VSAAHCYDDGAGYSWGDGAGSAGNAVGGT ADBGEWPWQVSLHALGGGHICGASLISWN VSAAHCYDDGAGYSWGGAGCAGRNKFG TGWHYGHTYGGAGAGAGAGAGAGAAACAACAACAACAACAACAACAACA	<del></del> /	1171	Λ.	3370	1202	1070	
NSPWRSQVLAAURAVGAGPWPPPPGSPLA   PSSRSAPGLASGPAPALIQOGWISSAGPLIC   NGTILALGIKP**AWGWGEWRKG     A							
PSSRSAPGLASGPAAPLLQGVHSSAGPLLC							
NGTLALGLEP**AWGWEWRPKG	1				1	ĺ	
448 1798 A 3604 3115 557 FREKGGGGFKDFGAGLKYNSRHEKVNGLI GVEFLPVINNYKVEKEHGREWVLAANL LLLVLLGIGFLVWHLQYRDVRVQKVFNGY RITNENFYDAYENSNSTEFVSLASKVKDAL LYSGVFFLGPYHKESAVTAFEGGSVIAYYW FSIPQHLVEEAERWMAEERVVMLPPRARSI FVVTSVVAPTDSKTVQHTYNSCSFGI-H, GVELMRFTTFGFPDSPYPAHARCQWALKG ADSVLSLTFRSFDLASCDGRHLWTYVTN SPMEPHAU VQLCGTYPPSYNLTFHSSQNI LITLITNTERRHPGFEATFFQLPRMSSCGG RKAQGTINSPYYPGHFYDDICTVMIEVYM QHYKVRFKFFYLLEFGVPAGTCPFKDYED EKYCGERSGVVTYSNSNITVRFHSDQSYT GFLAEYLSYDSSDPCPGGFTCRTGRCIKE CDGWADCTDHSDELINGSAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEASK KVNVYTCTKHTYRCLNGLLSKGNPFCDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHLCGASLISFN VSAAHCYDDRGFRYSDPTQWTAFLGLHD QRSAPGVGERLKRISHPFNDFTFDVDLA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQVGGTGALLQKGEIRVINQTTCI LLPQQTTRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPG TTLPLFFDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWS EMTRPSLMAGRGHIGWSAQQSATVANPY ANPDLLPHFLGEPEDVYIVKNEPULVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSC TMEVRINVSRQQVEKVPGLEEVWCQCVA SSGTTKSQKAYIRIAYLRKNFEQELAKEV CQGIVLPCPPEGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASWS ADAAEVFGTKGHSQQGAMPHIPEDEEPP PQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGETDKNLANKVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNHQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSFGDSP VSSNLSPSSAFTSSRSNSLTPTVSRTPFEGGEAD SPHPGEFNVPKGLADRKQNDQRKVSQGRI RPPPVEKSEKELIEGCKENDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK WEGSPLPRSTTQDAAGVGPGPAGQGGPAG MGPEAGSKAELPFTVSRTPLRGLSWDSGI RPPPVEKSEKELIEGCKENDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK WEGSPLPSSTTODAAGVGPPASQGGGPAG MGPEAGSKAELPFTVSRTPLRGLSWDSGI RPPPVEKSELIEGCKENDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK WEGSPLPSSTTODAAGVGPPASQGGGPAG MGPEAGSKAELPFTVSRTPLRGLSWDSGI RPPPVEKSELAIEGCKENDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKAGLK APGLKDFQUQVPYKMQKTKKLREEHLM APGLKDFQUQVPYKMQKTKKLREEHLM	ł					l .	
GVEPLPVNNVKKVEKHOPGRWVVLAAVL LLLVLLGIGFLVWHLQYRDVYQKVPNGY RITNENFVDAYENSNSTEFVSLASKVKDAL LYSGVPFLGPYHKESAVTAFSEGSVIAYVW FSISPQHLVEEARVMAERERVVMLPPRARSI FVVTSVVAPFDSKTVQRTONSCSGGLH GVELMRPTTGPPDSPYPAHARCQWALRG ADSVLSLTRESPDLASCDERGRHLVVTVN SPMEPHALVQLCGTYPPSYNLTHSISQNY LITLITNTERRHGFEATFFQLPRMSSCGGG RKAQGTRNSPYYPGHYPPNIDCTVMIEVPN QHVKVRFKFFYLLEFGVAGTCPKDYVED EKYCGERSQFVYTSNSNKITVRFHSDQSYT GFLAEVLSYDSSDPCPGCTGTGCRKEE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEASS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSBEDCDCCGLRSTRQARVVGGT ADEGEWPWQVSLHALQQGHICGASLISPN VSAAHCYDDRGFRYSDPVTAFLGLID QRSAPGVQERRLKRIISHPFRDFTFDYDLA ELEKPAEYSSMYPHICLPDASHVFPAGKAT TGWGHTQYGGTGALLQKGEIRVINQTTCL LLPQQITPRMMCVGFLSGGVDSCQGDSG SSVEADGRIFQAGVVSWGDGCAQRNKPG TRLPLFRDWKENTGV SSVEADGRIFQAGVVSWGDGCAQRNKPG TRLPLFRDWKENTGV ANPDLLPHFLGFPEDVYIVKNKPVLLVCK PATQIFFRCNGEWVRQVDHYRSTDGSSS TMEVRINVSRQQVEKVFGLEEVWCQCVA SSGTTKSQRAYIRLAYLRKFSGPLAKEV ANPDLLPHFLGFPEDVYIVKNKPVLLVCK PATQIFFRCNGEWVRQVDHYRSTDGSSS TMEVRINVSRQQVEKVFGLEEVWCQCVA SSGTTKSQRAYIRLAYLRKFSPELKEV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSFGSPF VSNNLSPSASFTSSRSNSLTYTFPEGDEAD SPHPGEFNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKERLIEQKENDPLQYPETTPLG PVTNSSGKMALNSPQFGPVESELGKQLLK WEGSPLPRSFTQDAAGVGRAGGGEL WEGSPLPRSFTQDAAGVGNQGRAGGGL RPPPVEKSKERLIEQKENDPLQYPETTPLG PVTNSSGKMALNSPQFGPVESELGKQLLK WEGSPLPRSFTQDAAGVGNAGGRAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGRAGG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL WEGSPLPRSFTQDAAGVGNAGGRAGGRAGG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQQVQVYNMQKLTKLREEHLLM APGLKDFQQQQVYPTMQKLTKLREEHLLM APGLKDFQQQQVYVMCKLTKLREEHLLM APGLKDFQQQQVYTGARGTAGGRAGGKL APGLKDFQQQQVYTGCRAGGRAGGKL						L	NGTLALGLKP**AWGWGEWRPKG
LLI-VILIGIGFLVWHI-QYRDVRYQKYNOG RITTNENFVDA YENNSITEFYSIL ASKVEDAL LYSGVPFLGPYHKESA VTAFSEGSVIA YYW FSIPQHI-VEEAERWMAEERVYMLPPRARSI FVVTSVVAPTDSKTVQHENSCSFGIL-WITVYN SPMEPHAL VQL.GTYPPSYNLTHISSQNI LITILTNTERSPIDLASCOGG RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHYKVRFKFFYLLERGVPAGTCPFLGYYEN EKYCGERSGVVTYSNNSHITVRFHSDQSYT GFLAEVI-SYDSSDPCFGGFTCRTGRCIRKE CDGWADCTDHSDELINGSAGGG RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHYKVRFKFFYLLERGVPAGTCPFLGYYEN CKPLFWVCDSLNDCGDNSDEGGCSCPAQ RCSNGKCLSKSQQCNGKAGHOFTCKN CKPLFWVCDSLNDCGDNSDEGGCSCPAQ RCSNGKCLSKSQQCNGKOGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLSSFTRQARVVGGT ADEGEWPWQVSLHALGGGHILGASLISFN VSAAHCYDDRGFRYSDPTQWTAFLGLID QRSAPGVGERLKRIISHPFNDFTFDYDLA ELEKPAFYSSMYRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLQGGHILGASLISFN VSAAHCYDDRGFRYSDPTQWTAFLGLID QRSAPGVGERLKRIISHPFNDFTFDYDLA ELEKPAFYSSMYRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLQGKGRANGYNSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGT TLPPFRDWIKENTICS LIPQGTPRMCVGPLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGT TRPLFIRDWIKENTICS  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWS EMTRRPSLMAGRQHGWSQQSATVANVY ANPDLLPHFLGEPEDVYIVNKPVLLVCK, PATQIFFK CNGGEWNQQVIPKRSTDGSS TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV PATQIFFK CNGGEWNQQVIPKRSTDGSS TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSGKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSGKARSSFTSSNSNSTSSSSSSSSSSSSSSSSSSSSSSSSSSSS	448	1798	Α	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE
LLI-VILIGIGFLVWHI-QYRDVRYQKYNOG RITTNENFVDA YENNSITEFYSIL ASKVEDAL LYSGVPFLGPYHKESA VTAFSEGSVIA YYW FSIPQHI-VEEAERWMAEERVYMLPPRARSI FVVTSVVAPTDSKTVQHENSCSFGIL-WITVYN SPMEPHAL VQL.GTYPPSYNLTHISSQNI LITILTNTERSPIDLASCOGG RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHYKVRFKFFYLLERGVPAGTCPFLGYYEN EKYCGERSGVVTYSNNSHITVRFHSDQSYT GFLAEVI-SYDSSDPCFGGFTCRTGRCIRKE CDGWADCTDHSDELINGSAGGG RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHYKVRFKFFYLLERGVPAGTCPFLGYYEN CKPLFWVCDSLNDCGDNSDEGGCSCPAQ RCSNGKCLSKSQQCNGKAGHOFTCKN CKPLFWVCDSLNDCGDNSDEGGCSCPAQ RCSNGKCLSKSQQCNGKOGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLSSFTRQARVVGGT ADEGEWPWQVSLHALGGGHILGASLISFN VSAAHCYDDRGFRYSDPTQWTAFLGLID QRSAPGVGERLKRIISHPFNDFTFDYDLA ELEKPAFYSSMYRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLQGGHILGASLISFN VSAAHCYDDRGFRYSDPTQWTAFLGLID QRSAPGVGERLKRIISHPFNDFTFDYDLA ELEKPAFYSSMYRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLQGKGRANGYNSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGT TLPPFRDWIKENTICS LIPQGTPRMCVGPLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGT TRPLFIRDWIKENTICS  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWS EMTRRPSLMAGRQHGWSQQSATVANVY ANPDLLPHFLGEPEDVYIVNKPVLLVCK, PATQIFFK CNGGEWNQQVIPKRSTDGSS TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV PATQIFFK CNGGEWNQQVIPKRSTDGSS TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSGKAYIRIAYLRKNFEQEPLAKEV CGGTUPLCPFGGIPPAG  450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASSI TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSGKARSSFTSSNSNSTSSSSSSSSSSSSSSSSSSSSSSSSSSSS	1		•			ŀ	GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG
RITNENFVDAYENSNTEFVSLSKVLDAU LYSGVPPLGPYHKESAVTAFSEGSVIAYYM FSIPQHILVEEAERVMAEERVVMLPPRARSI FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAV GVELMRPTTPGFPSPSPYAHARCQWALRG ADSVLSLTFRSFDLASCDERGRILVTVYN SPMEPHAU VQLGGTYPSYNLTFISSSQN LITLITNTERRHPGFEATFFOLPRMSSCGG RKAQGTFNSPYYPGHYPPNIDCTWNIEVPF QHVKVYFKFFYLLEFQFVAGTCFKDYVEB EKYCGERSGFVVTSNSNKITVRFISDQSYT GFLAEVILSYDSSDPFCGFCRGRCIKRE CDGWADCTDHSDELNCSCDAGHOFTCKN CKPLFWVCDSLNDCGDNSDEGGGSCPKQ RCSNGKCLSKSQQCNGKDDCGDGSDEASG KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISFN VSAAHCYDDRGFRYSDTQWTAFLGLHD QRSAPGVOERRLKRIISHPFNDFTDVTDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLQKGERVNOTTCI LLPQQITPRMMCVGFLSGGVDSCQDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGT TRLPLFRDWIKENTGV ANPDLLPHFLGEFDVYTVKNKPLLVCK, PATQIFFKCNGEWVRQVHVIERSTDGSSC TMEVRINVSRQVEKVFGLEEFWCQCVA SGGTTKSQKAYRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQGCAAMPHIPEDEEPP PQAAQSPAGQQGPTAGVSCSPTTTVLTG STPSGETDKNLARRYNSKSTSPGDSP VSPNLFSSASPTSSKSNSLTVPTPFEGDEAD SPHFGETNVFKGLADRKONDGKVSQGGR AGPAGSKAELPPTVSRPPLLRGLSWDSGI PGFRLQKVLAKLPLAEEEKRARGKGGKL APGLKDFQIQVQPVGQPLAYEGRBADG PFRLQKVLAKLPLAEEEKRARGKGGGRAA MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGFRLQKVLAKLPLAEEEKRARGKGGGRAA					1		
LYSGYPFLGPYHKESAVTAFSEGSVIAYYW FSIPQHLVEEAERVMAERVVMLPPRARSI FVYTSVVAFPTDSKTVQRTQDNSCSFGLHA GVELMRPTTPGFPDSPYPAHARCQWALRG ADSVLSLTRSFDLASCBGRHLVNTYN SPMEPHAULVQLGGTYPPSYNLTFHSSQNI LITLITNTERRHPGFEATFFQLPRMSSCGGF RKAQGTFNSPYYPGHPYPNIDCTWNIEVPN QHVKVRFRFFYLLEFGVPAGTCFRDYVEED EKYCGERSGFVYTSNSHITVRFHSDQSYT GFLAEYLSYDSSDPCPQGFTCRTGRCRKEE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEAS KVNVVTCHTTYRCHOLCLSKGNPECD EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEPWQVSLHACTGNGLCLSKGNPECD EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEPWQVSLHACTGNGLLOKGGIRVNOTTCI LLPQQITPRMMCVGFTCLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGGIRVNQVTCI LLPQQITPRMMCVGFTGSQFDSCQDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV ANDLLPHFLGPEDVIVKNKPVLLVCK PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQVEKVFGLESP WCCCVA SGGTTKSQKAYRIAYLRKNFPQGFLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTGPGVSFACGAQASWST ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSFAGQQGPTTAGVSCSPITTIVLTG STPEGGETDKNLANRVHSHKRLSSRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGKSRSSPGDSP VSPNLSPSASPTSSRNSLTVPTPPFGGEAD SPHGETNVFKGLADRKQNDGKVSQGGLAG SPHGGENVFKGLADRKONDGKVSQGGRAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGFRLQKVLAKLPLAEEEKRRAKSGGGRAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGFRLQKVLAKLPLAEEEKRRAKGAGGKL APGLKDPGIQVQPVMQKLTKLREEPHLM					ļ		
FSIPQHILVEEAERVMAEERVVMLPPRARSI FVYTSVVAFPTDSKTVQRTQDNSCSFGLH, GVELMRPTTPGFPDSPYPAHARQWALRG ADSVLSLTFRSFDLASCDERGRHLVTTYN SPMEPHAL VQLCGTYPSYNLTFISSIQN LITLITINTERRHPGFEATFFOLPRMSSCGG RKAQGTPNSPYYPGHYPPNIDCTWNIEVPR QHYKVYFRFFYLLEFQPVAGTCPRCHYVEB EKYCGERSQFVVTSNSNRITVRFHSQSYT GFLAEYLSYDSSDPCGFTCRTGRCIRKE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEGCSCPLAG RCSNGKCLSKSQQCMSKDDCGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECD EDCSDGSDEKDCDCGASFTRQARVVGGT ADEGEWPWQVSLHALQGGHICGASLISPN VSAAHCYIDDRGFRYSDFTQWTAFLGLHD QRSAPGVQERLKRISHFFNDFTFDYDLA ELEKPAEYSSMVRPICLPDASHTPAGKAT TGWGHTQYGGTGALLQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGT TRLPLFRGWIKENTGO ANDELPHFLGEFEDVTYNKNRPVLLVCK PATQUFFKCNGGWVRQVDHVIERSTDGSSC TMEVRINNSQQVEKVGLESTWCQCVA SSGTTKSQKAVIRAYLRKNFEQEPLAKEV EQGIVLPCRPPBGIPPAE  450 1800 A 3620 1 2676 MEPSLGGMDLTCPFGVSFACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPP PQAAGSPAGQGGPTAGVSCSFFTTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTHRGKKSRSSFGDSP VSPMLSPSASPTSSRSNSLTVPTPFFGDEAA SPHGGERDKALNSPQCPCASCGRPAG MGPEAGSKAALDPTVSRPPLLRGLSWDSGI PGPPTVEKSKELAIEQKENFDPLQYPETTFRG PVTNSSGKMALNSPQCPOPASQGRGPAG MGPEAGSKAALDPTVSRPPLLRGLSWDSGI PGPRLQKVLALPLAEEEKRAGKAGGKL APGLKDFQIQQVPYMQKLTKLREEBHLM	. ]			]	}	]	
FVVTSVVAFPTDSKTVQRTQDNSCSFGLLAL GVELMRFTTPGFPDSPYPAHARCQWALRG ADSVLSLTFRSFDLASCDERGRILVITYN SPMEPHAU VQLCGTYPPSYNLTFISSSQM LITLITNTERRHPGFEATFFQLPRMSSCGGG RKAQGTFNSPYYPGHYPPNIDCTWNIEVPN QHVKVRFKFYLLEPGVPAGTCPKDYVEPN QHVKVRFKFYLLEPGVPAGTCPKDYVEPN QHVKVRFKFYLLEPGVPAGTCPKDYVEPN QHVKVRFKFYLLEPGVPAGTCPKDYVEPN QHVKVTSNSNKITVRFHSDQSYT GFLAEYLSYDSSDPCPGQFTCRTGRCIKKEE CDGWADACTDHSDELNGCDAGHGTCKN CKPLFWVCDSLNDCGDNSDEQGGSCPAQ/ RCSNGKCLSKSQQCNGKDDCGDGSDEASG KVNVVTCTKHTYRCLOGLLSKGNPECDC EDCSDGSDEKDCDCGLRSTTRQARVVGGT ADEGEPWQVSLHALGGGHICGASLISFN VSAAHCYIDDRGFRYSDPTQWTAFLGLIBD QRSAPGVQRRLKRIISHPFNDFTFDVDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALLLQKGEIRVNQTTCI LLPQQTTPRMMCVOFFLSGGVDSCQGDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGV TRLPLFRMMCVAFLSGGVDSCQGDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGV TRLPLFRLGMENTGV ANPDLLPHFLGGEPDVIVNNKPVLLVCK PATQIFFKCNGGWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYRIAYLRKNFPQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSFACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSPAGQQGPTTAGVSCSPTTIVLTG STPEGETDKNLANRVHSHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSSFGDSP VSPNLSPSAASTSSRNSLTVPTTPFGGDEAD SPHGEPNVPKGLADRKQNDQRKVSQGRAA SPHPGEPNVPKGLADRKQNDQRKVSQGRAA SPHPGEPNVPKGLADRKQNDQRKVSQGRAA SPHPGEPNVPKGLADRKQNDQRKVSQGRAA MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRRAGKAGGKL APGLKDPGIQVQPVMQKLTKLREEPHLM	ļ						1
GVELMRPTTPGFPDSPYPAHARCQWALIS ADSVLSLTRRSPDLASCDERGRHLVITVYN SPMEPHALVQLCSCTYPPSYNLTFISISQN LITLITNTERRHPGFEATFFQLPRMSSCGGG RAQGTFNSPYYPGHYPPNIDCTWNIEVPN QHYKVRFKFFYLLEPGVPAGTCPKDYVEDE EKYCGERSQFVVTSNSNKITVRPHSDQSYT GFLAEYLSYDSSDPCPGQFTCRTGRCRKEKE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLPWVOSLNDGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEASC KVNVYTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSBERDCDGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISFN VSAAHCYDDDRGFSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVPPAGKAT TOWGHTQYGGTGALLQKGEIRVNQTTCI LLPQQTTPRMMCVQFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWKENTGW  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRSIMAGRQHGWSAQQSATVANPY ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQVEKVFGLEEYWCQCVA' SSGTTKSQKAYIRIAYLRKNFFQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHPEDEEPPP PQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDFAGHIDLVNDQLFDISISEEDKK LALLEEAKLVSERPLTRGRKSRSSPGDSP VSPNLSPASSFTSSRSNSLTVPTPPFGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPFVEKSKELAIEQKENFPPLQYPETIPKG PVTNSSGKMALNSPQPGPVSELGKQLLK' WGGSPLPSSPTODAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI RGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDPQQVQVPRMQKLTKLREEHILM	. 1		1		1		
ADSVLSLTERSFDLASCDERGRIL VITVYN SPMEPHALLVQLCGTYPPSYNLTFHSISQN LITLITNTERRHPGIFEATFFQLPRMSSCGG RKAQGTFNSPYYGHYPPNIDCTWNIEVPN QHYKVPRKFFYLLEPGYPAGTCPKDYVED EKYCGERSQFVYTSNSNKITVRFHSDQSYT GFLAEYLSYDSSDPCPGQFTCKTRGCREKE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ; RCSNGKCLSKSQQCNGKDDCGDSSDEAS KVNVYTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYUDDRGFYSDPTQWTAFLGLHD QRSAPGVQERRLKRIIGSIPFNDFTFDYDLA ELEKPATYSSMYMPICLPDASSHYPAGKAT TGWGHTQYGGTGALILQKGEIRVINQTTC! LLPQGTTPRMCVGFLSGGVDSCQGDSGG SSVEADGRFQAGVVSWGDGCAQRNKPGT TRLPLFRDWKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNRFVLLVCK, PATQIFFKCNGEWQDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVA' SSGTTKSQKAYIRAYLKNFFQEPLAKEV EQGIVLPCRPPEGIPTAGT FQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGGTDKNLANRVHSPHKRLSHRHLKY ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSFGDSP VSPNLSPASSFTSSRSNSLTVPTPFEGDEAD SPHPGEFNVFGLARKVODGRKVSQGRI RPPVEKSKELAIEQKENFDPLQYPETIFKG PVTNNSGKMALNSPQPGPVSELIGKQLK' WEGSPLRSPTODAAGVGPPASQGRGFAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKFRAGKAGGKL AFGLKDFQQVQVYRMQKLTKLREEHILM				}	ł		
SPMEPHAL VQLCGTYPPSYNLTHISISQON LITLITNTERRHPGIFATFFQLPRMSSCGG RKAQGIFNSPYYPGHYPPNIDCTWNIEVPN OHYKVRFKFFYLLEPGYPAGTCPKDYVEB EKYCGERSOFVYTSNSNKITVRPHSDQSYT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKEI CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLRSFTRQARVGGT ADEGEWPWQVSLHALGQGHICGASLISFN VSAAHCYUDDGFSSPDFTQWTAFLGLHD QRSAPGVQERRLKRIISHPFNDFTFDYDLA ELEKPAEYSSMVRPICLPDASHVFPAGKAI TGWGHTQYGGTGALLQKGEIRVNQTTCI LLPQQTFRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVBSCQCGDSGG SSVEADGRIFQAGVBSCQCGDSGG SSVEADGRIFQAGVMSQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVA SSGTTKSQKAYIRIAYLKRFQEPLAKEV EQGVIUPCRPPEGIPAE  450 1800 A 3620 1 2676 MEPSLQQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHPEDEEFP PQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGETUKNLANRVHSPHKRLSHRHLKV ASLTSVDFAGHIIDLVNDQLDFISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSFGDSP VSPNLSPASSFTSSRSNSLTVPTPPEGDEAD SPHPGEFNVPKGLADRXQNDGRKVSQGRI RPPVVEKSKEALIEQKENFDPLQYPETTPKG PVTNSSGAMALNSPQPGPVSEGLGKQLLK' WEGSPLPSSTTODAAGVOPPASQGRGFAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI GPRLQKVLAKLPLAEEEKFFAGKAGGKL AFGLKDFQLVQVYRMQKLTKLREEHILM	Ì		•		1		
LITILITNTERRÄPGERATFOLPRMSSCÄGE RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHVKVRFKFFYLLEFGVPAGTCPKDYVEID EKYCGERSQFVVTSNSNKITVRPHSDQSYT GFLÆYLSYDSSDPCPGQFTCRTGRCIRKEI CDGWADCTDHSDELINGSCDAGHQFTCKN CKPLFWVCDSLINDCGBDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGGSSDEASQ RCSNGKCLSKSQQCNGKDDCGGSSDEASQ RCSNGKCLSKSQQCNGKDDCGGSSDEASQ KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSBEKDCDCGLRSTTRQARVVGGT ADEGEWPWQVSLHALGQGFICGASLISPN VSAAHCYTDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVPFICLPDASHVFPAGKAT TGWGHTQYGGTGALLLQKGEIRVNQTTCI LLPQQITPRMMCVGFLSGGVDSQQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYTRIAYLRKNFEQEPLAKEV EQQIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDORKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK: WEGSPLPRSTTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKKFAGKAGKIL APGLKDPQIQVQVYMQKLTKLREEHLIM			1	ļ	1	!	1
RKAQGTPNSPYYPGHYPPNIDCTWNIEVPN QHVKVRFKFYLLEPGVPAGTCPKDYVEN EKYCGERSQFVYTSNSNKITVRFHSDQSYT GFLAEYLSYDSSDPCPGOFTCRTGRCIRKE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ/ RCSNGKCLSKSQQCNGKDDCGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDGLRSTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYDDDRFYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFNDFTFDYDLA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGERIEVNQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWDGCGAQRNKPGT TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSILMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVNKRYVLLVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFQLEYYWCQCVA SSGTTKSQKAVRIAYLRKNFEQEPLAKEV EQGVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGGMDLTCPFGVSPACGAQASWSI ADAAAEVPGTRGHSQQEAAMPHIPEDEEPP PQAAQSPAGQQGPTTAGVSCSPTPTTVLTG TSPEGETDKNLANRYBKQLSAMPHIPEDEEPP PQAAQSPAGQQGPTTAGVSCSPTPTTVLTG TSPEGETDKNLANRYBKQLSAMPHIPEDEEPP VQAAQSPAGQQGPTTAGVSCSPTPTTVLTG TSPEGETDKNLANRYBKRLSTRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGEAD SPHPGEFNVFKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK: WEGSPLRSFTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPPLQKVLAKLPLAEEEKKFAGKAGGKL APGLKDPQIQVQVYMQKLTKLREEHLIM				1			SPMEPHA/LVQLCGTYPPSYNLTFHS/S/QNVL
OHYKVRFEFYLLEGVPAGTCPKDYVEB EKYCGERSQFVVTSNSNKITVRFHSDQSYT GFLAEYLSYDSSDPCPGGFTCRTGRCRKEE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ' RCSNGKCLSKSQCNGKDDCGDGSDEASS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCGGLSSTFRQARVGGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTDPJDLA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWCHTTQYGGTGALLLQKGEIRVNGTTCI LLPQQTTPRMWCVGFLSGGVDSCQGDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVYGLEEYWCQCVAY SSGTTKSQKAVRIAYLRKNFEQEPLAKEV EQGIVLPCRPPBGIPPAE  450 1800 A 3620 1 2676 MEPSLGGMDLTCPFGVSPACGAQASWSI ADAAAPVGTRGBOQEAAMPHIPEDEEPPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDGRVSQGRR SPPVEKSKELAIEQKENFOPLQYPETTPKG PVTNSSGKMALNSPQRPVESSELGKQLLK: WEGSPLRSSTYDDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPPLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDPQIQVQPVRMQKLTKLREEHLMW	i	•			ĺ	į	LITLITNTERRHPG\FEATFFQLPRMSSCGGRL
EKYCGERSOFVVTSNSNKITVRFH5DQSYT GFLAEYLSYDSSDPCPGGFTCRGCRIKEL CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ; RCSNGKCLSKSQQCNGKDDCGDGSDEAS; KVNVYTCTKHTYRCLNGLCLSKGNPECDG EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDFTQWTAFLGLHD QRSAPGVQERRLKRLISHPFPDDFTDVDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGEIRVNOTTCI LLPQQTTPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGFPEDVYIVKNKPVLLVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPY PQAAQSPAGQQGPFTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSPGDSP, VSPNLSPASPTSSRSNSLTVPTPPEGDEAD SPHPGERNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKEIAIEQKENFDPLQYPETTFKG PVTNSSGKMALNSPQPGPVESELGKQLLK: WEGSPLPRSSPTODAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKFRAGKAGGKI APGLKDFQUQVPVRMQKLTKLREEHILM	·						RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN
EKYCGERSOFVVTSNSNKITVRFH5DQSYT GFLAEYLSYDSSDPCPGGFTCRGCRIKEL CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ; RCSNGKCLSKSQQCNGKDDCGDGSDEAS; KVNVYTCTKHTYRCLNGLCLSKGNPECDG EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDFTQWTAFLGLHD QRSAPGVQERRLKRLISHPFPDDFTDVDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGEIRVNOTTCI LLPQQTTPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGFPEDVYIVKNKPVLLVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPY PQAAQSPAGQQGPFTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSPGDSP, VSPNLSPASPTSSRSNSLTVPTPPEGDEAD SPHPGERNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKEIAIEQKENFDPLQYPETTFKG PVTNSSGKMALNSPQPGPVESELGKQLLK: WEGSPLPRSSPTODAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKFRAGKAGGKI APGLKDFQUQVPVRMQKLTKLREEHILM			J		] -		OHVKVRFKFFYLLEPGVPAGTCPKDYVEING
GFLAEYLSYDSSDPCFGGFTCRTGRCIRKE CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCPAQ RCSNGKCLSKSQQCNGKDDCGDGSDEASC KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDFTQWTAFLGLHD QRSAPGVQERLKRIISHPFPDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGEIRVNOTTCI LLPQQTTPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGV TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEFEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQVEKVFGLEEYWCQCVA/ SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPTAGVSCSPTTPTULTG TSPEGETDKNLANRVHSPHKILSHRHLK/ ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSPGDSP/ VSPNLSPSASFTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPTVEKSKEIAIEQKENTPDLQYPETTTKG PVTNSSGKMALNSPQOFPVESELGKQLLK/ WEGSPLPRSPTQDAAGVGPPASQGROPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSG GPPALQKVLAKLPLAEEEKFRAGKAGGKL APGLKDFQUQVPVRMQKLTKLREEHLM						1	
CDGWADCTDHSDELNCSCDAGHQFTCKN CKPLFWVCDSLNDCGDNSDEQGCSCP\AQ' RCSNGKCLSKSQQCNGKDDCGDGSDEAS KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHLGASLISPN VSAAHCYIDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVPPAGKAN TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGT TRLPLFRDWKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEFWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVNSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEAKLVSERFLTRRGRKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALINSPQPGPVESSLGKQLLK' WEGSPLPRSFTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDPIQVQPVRMQKLTKLREEHILM					ļ		
CKPLFWVCDSLNDCGDNSDEQGCSCPIAQY RCSNGKCLSKSQQCNGKDDCGDGSDEASY KVNVVTCTKHTYTYRCLNGLCLSKGMPECDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQDDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPDVYIVKNKFVLLVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYTRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAANPHPEDEEPPY PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSV VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETITKG PVTNSSGKMALNSPQPGPVESELGKQLLK' WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDPQIQVQPVRMQKLTKLREEHILM			1	ļ	1	ł	
RCSNGKCLSKSQQCNGKDDGDGSDEASK KVNVVTTKHTYRCLNGLCLSKGNPECDC EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAT TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRFQAGVVSWGDGCAQRNKPG' TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQVEKVPGLEEYWCQCVAN SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPC PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHHIDLVNDQLPDISISEDKK LALLEEALVSERPLTRRGRKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK' WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			}	•	ļ	ì	
KVNVVTCTKHTYRCLNGLCLSKGNPECDC EDCSDGSBCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVPPAGKAFI TGWGHTQYGGTGALLLQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGV TRLPLFRDWIKENTIGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGFEPEDVYIVKNKPVLLVCKX PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRLAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAMPHIPEDEEPPY PQAAQSPAGQQGPPTAGVSCSPTTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPY VSPNLSFSASFTSSRSNSLTVPTPFEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKV WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEKKFAGKAGGKL APGLKDFQLQVVPVRMQKLTKLREEHILM			1	l		ļ	
EDCSDGSDEKDCDCGLRSFTRQARVVGGT ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAFYSSMVRPICLPDASHVFPAGKAF TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGV TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSIMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAT SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV' EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSFRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPV VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEFNVPKGLADRKQNDQRKVSQGRI RPPVVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK' WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI POPTLQKVLAKLPLAEEKEFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM	·		ĺ	l		Ì	
ADEGEWPWQVSLHALGQGHICGASLISPN VSAAHCYIDDRGFRYSDPTQWTAFLGLHD QRSAPGVQERRLKRIISHPFNDFTFDYDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAI' TGWGHTQYGGTGALILQKGEIRVINQTTC! LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/A PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSFACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPFTAGVSCSPIPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKY WEGSPLPRSFTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			i	1	1	į.	
VSAAHCYIDDRGFRYSDFTQWTAFLGHD QRSAPGYQERRLKRIISHPFFNDFIFDYDIA ELEKPAEYSSMVRPICLPDASHVFPAGKAI TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIPQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEFEDVYIVKNKPVLLVCK, PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSPAGQQGPPTAGVSCSPTPTTVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSFGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHFGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK: WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1			•	
QRSAPGVQERRLKRIISHPFFNDFTFDYDIA ELEKPAEYSSMVRPICLFDASHVFPAGKAIT TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGV TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPY PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASFTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ļ	<u> </u>		}	
ELEKPAEYSSMVRPICLPDASHVFPAGKAN TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVA' SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV- EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPPTAGVSCSPTPTTVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGERNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK' WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ļ	i			
TGWGHTQYGGTGALILQKGEIRVINQTTCI LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGV TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTIDGSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPC PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRI RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSFTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1		1		QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL
LLPQQITPRMMCVGFLSGGVDSCQGDSGG SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  TRLPLFRDWIKENTGV  FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGF PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM							ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV
SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGF PGPRLQKVLAKLPLAEEEKFFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM				)	Ì		TGWGHTQYGGTGALILQKGEIRVINQTTCEN
SSVEADGRIFQAGVVSWGDGCAQRNKPGY TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCK/ PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGF PGPRLQKVLAKLPLAEEEKFFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM				l			LLPOOITPRMMCVGFLSGGVDSCOGDSGGPL
TRLPLFRDWIKENTGV  449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAX SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLKI WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ļ	i			
449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSS EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPP PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			l	ł	1		,
EMTRRPSLMAGRQHGWSAQQSATVANPV ANPDLLPHFLGEPEDVYIVKNKPVLLVCKA PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPA PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGKSRSSPGDSPY VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM	440	1700	1	3618	2	612	<del></del>
ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAPATQIFFKCNGEWVRQVDHVIERSTDGSSCTMEVRINVSRQQVEKVFGLEEYWCQCVANSGTTKSQKAYIRIAYLRKNFEQEPLAKEVEQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIADAAEVPGTRGHSQQEAAMPHIPEDEEPPOPQAAQSPAGQQGPPTAGVSCSPTPTIVLTGGTSPEGETDKNLANRVHSPHKRLSHRHLKVASLTSVDPAGHIIDLVNDQLPDISISEEDKKLALLEEAKLVSERFLTRRGRKSRSSPGDSPSVSPNLSPSASPTSSRSNSLTVPTPPEGDEADSSPHPGEPNVPKGLADRKQNDQRKVSQGRIRPPVEKSKELAIEQKENFDPLQYPETTPKGPVTNSSGKMALNSPQPPVESELGKQLLKWEGSPLPRSPTQDAAGVGPPASQGRGPAGMGPEAGSKAELPPTVSRPPLLRGLSWDSGIPGPRLQKVLAKLPLAEEEKRFAGKAGGKLAPGLKDFQIQVQPVRMQKLTKLREEHILM	449	1/99	^	3019	1 2	013	
PATQIFFKCNGEWVRQVDHVIERSTDGSSC TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGKSRSSPGDSPS VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1	1			
TMEVRINVSRQQVEKVFGLEEYWCQCVAY SSGTTKSQKAYIRIAYLRKNFEQEPLAKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			[	Í		1	
SSGTTKSQKAYIRIAYLRKNFEQEPLÄKEV EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1	[	ľ		
EQGIVLPCRPPEGIPPAE  450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			J.	}	]	}	TMEVRINVSRQQVEKVFGLEEYWCQCVAWS
450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSI ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ļ	1	}	1	SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL
ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPO VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			]	1			EQGIVLPCRPPEGIPPAE
ADAAEVPGTRGHSQQEAAMPHIPEDEEPPO PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSPO VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKELAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGF PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM	450	1800	A	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG
PQAAQSPAGQQGPPTAGVSCSPTPTIVLTG TSPEGETDKNLANRVHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1	1	1	1	1
TSPEGETDKNLANR VHSPHKRLSHRHLKV ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKY WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM				1			
ASLTSVDPAGHIIDLVNDQLPDISISEEDKK LALLEEAKLVSERFLTRRGRKSRSSPGDSP VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKY WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM					1	1	1 ' ' '
LALLEEAKLVSERFLTRRĞRKSRSSPGDSP. VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLK. WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM		1	1	1	1		
VSPNLSPSASPTSSRSNSLTVPTPPEGDEAD SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKY WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM		1	l	l	1	1	
SPHPGEPNVPKGLADRKQNDQRKVSQGRL RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKY WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			1		1	1	L .
RPPPVEKSKEIAIEQKENFDPLQYPETTPKG PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM		}	1.	1	1		VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS
PVTNSSGKMALNSPQPGPVESELGKQLLKT WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ĺ	ſ	1	1	SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP
WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM		l	1	Į	1		RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA
WEGSPLPRSPTQDAAGVGPPASQGRGPAG MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			· ·	j .			PVTNSSGKMALNSPQPGPVESELGKQLLKTG
MGPEAGSKAELPPTVSRPPLLRGLSWDSGI PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM			ļ	l	J		WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP
PGPRLQKVLAKLPLAEEEKRFAGKAGGKL APGLKDFQIQVQPVRMQKLTKLREEHILM	'				į .		
APGLKDFQIQVQPVRMQKLTKLREEHILM		<u> </u>	!		ļ	1	
	}	1	1	l	1		
I I I I I I I I I I I I I I I I I I I		İ	1	l	1		
			1		1		QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE
3		1	Ì	1			ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP
		l	1	l	1	1	LTEKEVENVFVQLSSAFRNDSYTLESRINQAE
		ļ	1	1	1	1	RERNLTEENTEKELENFKASITSSASLWHHCE
HRETYQKLLEDIAVLHRLAARLSSRAEVV	L	<u></u>	1		1	· ·	HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide sequence	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVQCQHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWR\KGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT \HVGYSSSREITE\AAVLLFYR
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457	1807	A	3660	14	1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQP\S LNQEDIYITTESLT\TAAGSP\GTGEHVPGSEM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS SCGYVSTDQLNKIMP
458	1808	A	3663	154	462	TRAPASGRSGAGLALSANAPDSGGHPGATEG PAGSLAHASGSARGTWRVRGRGSHGWERTV GAGGCANPVPALHSCASAPRGTGRVSALGPK TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC \NQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM VLALNVVTVATITVRHFVNQRADYQ\YQKLQ NY
460	1810	A	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\ TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N CYYD/STKSFFYISCG*K\RKPTWAENRRLNA KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG HGS
463	1813	A	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRIRPKSSTK
464	1814	A	3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT FSTPSSPAPFPTSSNPSSATTPPNPSP\GQR\DSR FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR WHGEVAIRLLEMDGHNQDHLKLFKKEVMN YRQTRHENVVLFMGACMNPPHLAIITSFCKG RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF \(\text{GISGVVP\text{CISG}}\)CISGVVPLGRRENQLKLSHDWLCYLAPEIVR EMTPGKDEDQLPFSKAADVYAFGTVWYELQ ARDWPLKNQAAEASIWQIGSGEGMKRVLTS VSLGKEVSENLSACWAFDLQERPS\FSLLMD MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR FERFGLGVLESSNPKM
465	1815	A	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP QAPLLAVGISFGGILVLNHLAQARQAAGLVA ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

meleotide insertion  Sequence  sequence  sequence    Polycotytytalspritidalizpvlyt.saa	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
A66		7					nucleotide insertion FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
467 1817 A 3687 2465 837 ELPTPILAAHQLVNYVADHASSYHMKPIRMA RPGGPEHNEYALVSA WHSSAD DESGURHQ DDFDVSLLVCHCAAFFEEQGEAERHVLRLQF FVVLTSQRELFPRITADMRFRKPRPRIPPEPE APGSSAGSNEGSEAGULLAFPRAFLFPEPE APGSSAGSNEGSEAGULLAFPAFRLFPEPE APGSSAGSNEGSEAGULLAFPAFRLFILE PPCPPRILAEVG MARARLAQLVRLAGGICRRDTLWKRLFLILE PPCPPRILAEVG MARARLAQLVRLAGGICRRDTLWKRLFLILE PPCPPRILAEVG MARARLAQLVRLAGSPOLGT VLVINLOKFICVLVFLDSH LGKTSLTVVFREPFPVQPQDSESSPAQLVSTY HHLESVINTACTT WTELL GESQLPHMSLFL ESWAYQLACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQGEPSSNTHISH*PIQDSR ESGPRGPLGFFWGTFFGFPGRVSGVHTGWQ TPPRAPLESCPLIPLTTVSHLCPLSLRVTSHL DITAGHSHRDTWVPIPALPLKHLRPPSSPFA LGFWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQQRN QTCRKDSWPFFFFFLLCPLSLSSWNYRRPPFP GKQQRN QFFSQNDSAFFFFLFLLEFCSAAVTVQWH DFLSMQPPPPGFRQFTCLSILSSWNYRRPPPF PGNF*FLVKTGFPHVGGTGFELLTSSDLAPLA SQNGGTIGMSCAWFFFFFFFLCLC WAS SQNGGTGMSCAWFFFFFFFGLC MAYSWQTDFNPNSEREKQYENGEFLEVNQP HSSSQVSLGFDQIVDEISGRHYSEDENTTFY VPTAFK WDSTGISLNEAHQISLNEFTSKSREL SWHQYSKAPAGFSPSVLRYBEDENTTFY VPTAFK WDSTGISLNEAHQISLNEFTSKSREL SWHQYSKAPAGFSPSVLRYBEDENTFY VPTAFK WDSTGISLNEAHQISLNEFTSKSREL SWHQYSKAPAGFSPSVLRYBEDENTFY VPTAFK WDSTGISLNEAHQISLNEFTSKSREL SWHQYSKAPAGFSPSVLRYBEDENTFY VPTAFK WDSTGISLNEHENG WEBENEN YRHGFPSSIPPTNSFSSDFATHKEONG SPIGKHHGADDSRFSILAFSFTSLDKINLEKEL ENENHYHIGFESSIPPTNSSFSSDFACULTILRXY VDFHKKYKKIR ERYHAADVINNSGKIWSTTTAPFYQLFSKTK FNIHFIDNSTQPLHFMPCANYLVKDLAELIF FCTNDQLLFFKHILSVGSGCLTLIRKY VDFHKKYKKIR ERYHAADVINNSGKIWSTTAPFYQLFSKTK FNIHFIDNSTQPLHFMRCANYLVKDLAELIF TRYPQCFSKTK FNIHFIDNSTQPLHFMRCANYLVKDLAELIF TRYPQCFSTSKG LIEKVYTELSTSTYQLINYCKSFY ADFQPVNV PRCTSYLNPGLPSHLSTVYAAHNIEFWVHR NFPPLEKSLFRSMLTVKKFGACATINNAHLA AWTCLPLFFKEKSILGSMLFSMT,QSSPPVEM ITPGWDVSSSPSVPVT,QIDFATGWEMKPD SEENRSNLEEPLKECKHIARLSQKQTPLLISE EKKYLWFYFFYCVDNENGSIGACATN DEFENSERGKLIKLIGUIGERVKSASDHQRGVU LKNARBAYFKSWYQKLLAALQFCAGKAN DEFEKSECKLIKLIGDIGERVKSASDHQRGVU LKNARBAYFKSWYQKLLAALQFCAGKAN DEFEKSECKLIKLIGDIGERVKSASDHQRGVU	466	1816	A	3684	3	307	SSQYTVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
QCFHSQNDSAFFFLLLETEFCSAA/TVQWH DFLSMQPPPPGFKQFTCLSLLSSWNYRRPPPF PGNPI-FLVKTGPPHVQGTGFELLTSSDLAPLA SQNGGITGMSPCAWPFFFFFFGLC  469 1819 A 3714 4747 495 MAYSWQTDPNPNESHEKQYEHQEFLEVNQP HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPK WDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAFFTSLDKINLEKEL ENENHYHIGFESSIPPINSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTQENVYNIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTIELTSISYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSSEMHTILRRWTFSQPLEALGLLTSSPP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEQKLIKILGDIGERVKSASDHQRQEVL	467	1817	A	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAQLVRLAGGHCRRDTLWKRLFLLE PPGPDRLRLGGRLALAELEELLEAVHAKSIGD IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY HHLESVINTACFTLWTRLL*GSGLDH*MSLFL ESWAYQIACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ TPPRAPLPESCPL\PLTTVSHLCPLSLRVFTSHL DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG
469 1819 A 3714 4747 495 MAYSWQTDPNPNESHEKQYEHQEFLFVNQP HSSSQVSLGFDQIVDEISGKIPHYESEIDENTTF VPTAPK WDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNITNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENINYHIGFESSIPPTNSSFSSDFMYKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIINSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTQENVYNIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHLARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEPSKEQKLIKILGDIGERVKSASDHQRQEVL	468	1818	A	3691	960	499	QCFHSQNDSAFFFFLFLETEFCSAA/TVQWH DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF\*FLVKTGFPHVGQTGFELLTSSDLAPLA
2 Discontinuity of the control of th	469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENHNYHIGFESIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTQENVYNIIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					-	GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIIYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGGKNPQHFQDFVELCCRAYNIIR KHSQLLL\NLL\EMMLYAG\LPELSGI\QDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMILSNPIW
470	1820	A	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVQLLLSLGSVFLWALL
471	1821	Α	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STETI GGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN
473	1823	A	3746	3	500	*ISDKGLVSRICQELLRHLDAEQVSSTAGLSL THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753		5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP\ERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG LELATTFEHFYQHYMADRILSFGSSWLEGAV LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF QLQRLDKLFLEQEDEEEKRL*EEEEEEEEA EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG PHRRLQWTWLGRAELQFGKQILHVSTVQMW LLLKFNQTEEVSVETLLKDSDLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL GQGYVKRRDDRPQILMYAAPEPMGPCRGQA DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CPAQPTGAFIRAIVSSPEVISKYEKALRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH
	·					HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDEAYD
475	1825	A	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRVRASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	A	3763	267	1240	HILLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG PDEHPQDTDARDADGEAREREP/RRPSFAA*P VWGQPESPLPEASSAPPGPTLGTLPEVETIRA CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
<u></u>				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
479	1829	A	3766	2	2152	PSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRFILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR
480	1830	A	3777	251	3	FCEAMDNYFLMVPDHCTGLGLNC QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF VANFIMAKSWKQPTHPSVRT
481	1831	Α.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI
482	1832	A	3780	2	371	VLPVSCFQGQKFN TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP A*FFVFLVE\QGFTMLARMVSIS*PQ/CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

NO: of No		T GROTO	1 17-4	LOCO	D 31-4-3	F 50 19 4 1 1 1 1 1	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq- usnce of the corresponding to the c	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Aranine C=Cysteine,
uence USSN got Gersponding to last amino acid residue of peptide residue of peptide sequence peptide sequenc			noa	1 .			D=Aspartic Acid, E=Grutaniic Acid,
uence 9/496 correspondi ng to first am to eaid residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence per commendation of the period sequence per commendation of t		,					
uence since a since said residue of peptide sequence peptide sequence peptide sequence sequen	eotide	seq-	1	1	1		I=Isoleucine, K=Lysine, L=Leucine,
amino acid recidud of peptide sequence peptide sequence s	seq-	uence		1			M=Methionine, N=Asparagine, P=Proline,
residue of peptide   sequence	uence	j	J	914	ng to first		
residue of peptide   sequence			1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Peptide				l	residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
1838		ľ		1	peptide	٠ .	
1838			j			}	1 •
LCLLAFCLAGFSFVRGOVLFKGCDVKTTFVT HVPCTSCAALKGTCPSWIRELPDQITQDC YEVQLGGSWYSMSGCRKCKQWQKACC GYWGSKCHCCPGGAFTCCHGHCTLDGMD MGTCVCQENFRGSACQECQPTRRGPDCQSI CSCVHCVCNHGPRGDSCLCTAGVTGPHCD QELPVWQELGFPQNNFRLKAPNCKCLPG*P RNGLIATNPCRIP 489 1839 A 3822 934 669 FFFSEMESRSYTRLECSGAISAHLRLLGSNSIS ASAN VAGTIGACHHAQLFVTL-VETGFHHV QDGLDLLNIMHPPRPFWVLGFQA 490 1840 A 3825 79 9748 GCQSCVPAWRILRRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRGFASAGARLGRKAPV GLOGUNDARLAFRALLAHVLSON ASSPEASAGDCRAGRETAMATLEKLMKAPESS KSRQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPP FPPPPPQLPGPPPQAQPLLPQPPPPPPPPP FPPPPPQLPGPPPQAQPLLPQPPPPPPPPPP GPAVAEELHAPKELSATKCRWINKICHTC ENIVAQSVRINSPEFQKLLGJAMELFLLCSDDA ESDVRWADECLINKIKALMDSIN,PRLQLE YKEIKKNGAPRSLRAALWRFAELAHLVRRQD CRPYLVMLPCLTRTSKRPEESVGETLAAAV KIMASFGSFANDNEIKVLLKAFIANLKSSSTY RTAAGASAVSICHGHAKEAGGRSKAPV KEIKKNGAPRSLRAALWRFAELAHLVRRQD KIMASFGSFANDNEIKVLLKAFIANLKSSSTY RTAAGASAVSICHAAKESGGGRSKAPVLNLVLL LLYPVEDEHSTLLILGVLLTRYLVPLLQQU KDTSILKGSFGVTRKEMEVSPASAGLVQVYEE TLHHTTQHQDHNVTGALELQQLFRTPPPEL LQTLTAVGGGIQLTAAKEBGGGRSKQSVVYPDBAWA ATGGGGSSCSPVLSKKQGGKVLLGEEBALEDDD ESRSDVSSSALTASVEDISGELAASGVSTP SAGHDITTGCPRSQVTLAGEGARADLNDC TQASSPISDSSOTTTEGPDSAVTPSDSSEIVLJ GTDNQT/LGLQIGOPQDEDEATIGPDBASE FRNSMALQQAHLLKNMSHCRQPSDSSVDK KLRDEATERDGDENKPCRIGGDIGGSTDDDS APILVHCVRLLSASFLLTGGKNVLVPDRDVR SVKALALSCVGAAVALPESFFRKYLKYPLL TTEYPEGQYNDLHVIDHODPQVBGATALL GTLGSLSRSPRIVORDWGRTRILTGNTSL ADCTELLRKTIKDESSVTCLACTAVRNCVM SLCSSSVSELIG.UIDHVLTLRNSSVILVTE LETLAEIDFFLVSPLBAKABULHRCAHHYTO LKLQERVYNNVHLLGDEDFRVRHVAAAS RLVYKLFYRCDQGQADPVAVAVARDQSSVY KLLMHETQFSHSVSTITRIVRGYNLLSSTIT VIMENNLSVIAAVSHELITSTTRALTFGCC ALCLLSTAFPVCWSLGWHQCVPPLSASDSSV KLLMHETGPSHSVSTITRIVRGYNLLSHTIT VIMENNLSVIAAVSHELITSTRALTFGCC ALCLLSTAFPVCWSLGWHQCVPPLSASDSSV GRYNTISKSSSLGSFYHLPSYLKLHDVILKATH NYKVTILDLQNSTERFGGFRSASLAVATAK GKEEGEGASVFLYRWQEUGSPRISASGROST GPVTTSKSSSLGSFYHLPSYLKLHDVILKATH NYKVTILLQUNGTGGFGGFRSALVALVSQUE ATTQDIGKCVEELLCYHKSGFSREP	400	1929	<del> </del>	2818	<del></del>	701	
HIVPCTSCAAIKKGTCPSGWLRELPPDQTQDC YEVQLGGMVVSMSCGRRCKRCKQVQKACC GYWGSRCHECPGGAETFCNGHGTCLDGMDI NGTTCVCQENFEGSACQPCQDPNREGPDCQSI CSCVHGVCNHGPRGDGSCLCTAGVTGPHGD QELPVWQELGPPGMNRELKAPNCKCLPG** RNGLIATTNPCRP  489 1839 A 3822 934 669 FFSEMESRSVTILECSGAISAHLRLLGSSNSI ASAS*VAGTIGACHHAQLIFVTLVETGFHHV QDGLDLLNLMHPPRPFKVLOFQA  490 1840 A 3825 79 9748 GCQSCVPAWPRLRKRGPGSAGARLGRAPV GLGRGVQDGPLRCFYLRRAFHAVLSGA ASRPEASGDCRAGRETAMATLEKLIKAFES KSPQQQQQQQQQQQQQQQQQQQQQQQQQQQQP PPPPPPPP	400	1030	^	3616	] *	/61	I CLI A DCI A CECEVO COVI EV CODVETTEVT
YEVQLOGSMYSMSGCRIRCKQVVQKACC GYWGSRCHCCYGGAETTCCHGHOTLDGMDI NGTCVCQENFRGBACQECQDPNRFGPDCQSS CSCVHOVCNHGPRGDSGCLTAGVTGPHCD QELPVWQELGFPQNNPRLRKAPNCKCLPG*F RNGLIATPNPCRP RNGLIATPNPCRP 489 1839 A 3822 934 669 FFFSEMESRSYTRLECSGAISAHLRLLGSSNSI ASAS*VAGTIGACHHAQLIFVELVETGFHHV QDGLDLNIM.HIPPPRFVLGFQA 490 1840 A 3825 79 9748 GCQSCWPAWPRLRRRGPASAGARLGRKAPY GLPGWQDGRLBRCTYLRRAPFIAVLSGA ASRPASGDCRAGRETAMATLEKLMKAFFSS KSFQQQQQQQQQQQQQQQQQQQQQQPP PPPPPPPPQLPQPPPQAQPLLPQPPPPPPPPP GPAVAEELHRPKELSATKCRWNICLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRWADECLINSTKAPESVQETLAAAV KIMASFGNFANDNEIKYLLKAFANLKSSSPT RRTAAGSAVSICQHSRRTQYTSWLDNVLL LVPYEDBHSTLLIGUFLTRYVLQQQ KDTSLKGSFGTTKEMEVSPSABQLVQYYEI TLHHTTQHQDHNVTGALELLQQLERTPPEL LQTLTAVGGIQQLTAAXEESGGRSSGSIVEI AGGGSSSVPLSRKQGSVLLGEEALEDD ESRSDVSSALTASVKDEISGELAASSGVSTP SAGHDITTQPRSQHTQABSVDLASCDLTSS ATDGDEEDLSHSSSQVSAVPSDPAMDLNDG TQASSFISDSSGTTTEGPDSAVTPSSESIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE FRNSMALQQAHLKINNSHCRQPSDSSVDK VLRDEATEPGQDENKPCRRGDIGGSTDDDS APLVHCVRLLASSFLTGGKNULYDRDVRS SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGVYSDLMYDDHODPQVRGATAIL GTILCSLSRSPHYLOGDWATURDHODPQVRGATAIL GTLICSLSRSPHYLOGDWATURDHODPQVRGATAIL GTLICSLSRSPHYLOGDWATURDHODPQVRGATAIL GTLICSLSRSPHYLOGDWATURDHODPQVRGATAIL GTLICSLSRSPHYLOGDWATURLTGNTST ADCIPLLRTLTADESS VTCLACTAVRNCVM SLCSSSSVSSELGLQLIDVLTLRSSSYTWLPSTIT VTMENNLSRVIAAVSHELITSTTRALTFGCCI LKLQERVLNNVVHILLGDEDPRVHYAAASI RLVPKLFYKCDQGQADPVAVARDQSSVW KLMHETQPPSHPSVSTITHYRGYNJLSPITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI LKLQERVLNNVVHILLGDEDPRVHYAAASI RLVPKLFYKCDQGQADPVAVARDQSSSVK KLMHETQPPSHPSVSTITHYRGYNJLSPITT VTMENNLSRVIAAVSHELITSTTRALTFGCS ALCLSTAPPCVCRUSCHWATACSPISABDS KSCTVGMATMILTLLSSAWPLDLSAHDAVA QEEWPALGDRALVPMVEQLFSHLLKVNIK GREEGGASVPLSFRKGGEASAASRQSDT GPVTTSKSSSLGSPYHLPSYLKLIEDVILASTHE NYKVTJLQQNSTERGGFRSAAJSCSGSGSSVCSSSVFGGGANGACFSSCFREMMATVC VQCLLSTLFGTNLASGPGGETSSALDVSKGQGA QRLGSSSVRGILYHYCFMAPYTHFTQALAD LRNMVQAEQEADDTSGWPDULVQSTQLK  URQLSTAFFCCHSTARDSTGWPDULVSCTICK ATUQDISCVFTICKYSTGCK							
GYWGSRCHECPGGAETPCNGHGTLLOMDN				1			HVPC1SCAAIKKQ1CPSGWLRELPDQ11QDCR
NGTCVCQENFRGSACQEQOPNRFGFDHOO		l	1	i			YEVQLGGSMVSMSGCRRCRRQVVQRACCP
CSCVHGYCNIGPRGDÖSCLCFAGYTGPHONERLAPICKCLPG*				ł			
GELPYWOELGFPQNNPRIRKAPNCKCLPG*F   RNGLIATPRIPCRP		ł	l	ì		ĺ	NGTCVCQENFRGSACQECQDPNRFGPDCQSV
RNGLIATPHPCRP			i				CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
RNGLIATPHPCRP		1		Į.	1		OELPVWOELGFPQNNPRLRKAPNCKCLPG*H
489 1839 A 3822 934 669 FFFSEMESRSYTRLECSGAISAHLRILGSINSI ASSA'S-VAGTIGACHHAQLIFYFI-VETGFHHV QGGLDLLANLMIHPPRPKVLGFQA  490 1840 A 3825 79 9748 GCQSCWPAWPRLRRRPASAGARLGRKAPV GLPGRVQDGRPLRRCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRET-MATLEKLMKAFES KSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ						·	RNGLIATPNPCRP
ASAS*VAGTIGACHHAQLEFVFLVETGFHHV QDGDLLIANIMIPPRPREVIGE GA  490 1840 A 3825 79 9748 GCQSCWPAWPRLRRGPASAGARLGRKAPV GLPGRVQDGRPLRFCFYLRPAPTAPVLSGA ASRPASGDCRAGRETAMATLEKLMKAPFSS KSFQQQQQQQQQQQQQQQQQPPP PPPPPPPPQLPQPPPPPPPP	180	1830	Δ	3822	934	669	FFFSFMFSRSVTRLECSGAISAHLRLLGSSNSP
Opglollanimimprppkvl.grQa	407	1039	Α	3022	1 224	007	
1840   A   3825   79   9748   GCOSCWPAWPILERRGPASAGAILGRKAPY GLPGRVOUGGPLIPGCPY GENERATION OF GENERAL CONTROL OF GENERAL CONT		1	l	1	}		· ·
GLPGRVQDGRPLRFCFYLRPRAPIAPULSGA ASRPASGDCRAGRETAMATLEKLMKAFES KSFQQQQQQQQQQQQQQQQQQQPPP PPPPPPPPPPPPPPP				<del> </del>			
ASRPEASODCRAGRETAMATLEKLMKAFES KSFQQQQQQQQQQQQQQQQQQPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GRAVAEEPLHRKKELSATKKDVHICLTIC ENIVAQSVRNSPEFQKLLGIAMELFILCSDD/ ESDVRMYADECLNKVIKALMDSNLPRLQLE YKEIKKNGAPRSLRAALWRFAELAHLVRPQI CRPYLVNLLPCLTRISKRPESVQETLAAAV KMASFGNFANDNEIKVLLKAFIANLKSSSYT RRTAAGSAVSIQHSRKIYLKAFIANLKSSSYT RRTAAGSAVSIQHSRKYLVLKAFIANLKSSSYT RRTAAGSAVSIQHSRKYLVLKAFIANLKSSSYT RRTAAGSAVSIQHSRKYLVLKQHINVLL LLVPVEDEHSTLLILGYLTTRYTVPLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVYTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSGSIVEI AGGGSSSVLSRKQKGVLLGEEEALEDD ESRSDVSSSALTASVKDEISGELAASSOVSTPY SAGHDITTEQPRSOHTLAAKESGGRSGSIVEI AGGGSSSVLSRKQKGVLLGEEEALEDD ESRSDVSSSALTASVKDEISGELAASSOVSTPY SAGHDITTEQPRSOHTLAADSVDLAASCOLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSGTTTEGPDSAVTTSDSSEIVLD GTDNQYLGQIQQPQDEDEEATGILPPEASE. FRNSSMALQQAHLLKNMSHCRQSDSSVDK VLRDEATEGPGDENKPCRIGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGGVSDINLYDHDDPQVRGATALL GTLICSILSRSRPHVGDWMGTRITLTGNTFSL ACKPLKRTLKDESSVTUNDHODPQVRGATALL GTLICSILSRSRPHVGDWMGTRITLTGNTFSL ACKPLKTLKDESSVTLACTAVRNCVAL SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSRLEAKAENLHRGAHHYTG LKLQERVINNVHLLIGDEDRVRRIVAAAS IRLVPKLFYKCDGQADPVVAVAADOSSVY KLLMHETOPPSHFSVSTTRIVRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPRGLFSHLLKVINI AHVLDDVAPGPAIKAALPSLTNPSLSPIRRK GKEKEPGGASVPLSPKKGSBEANPAATK QEEVWPALGDRALVPALGPSLFSHLKVINI AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGGASVPLSPKKGSERSENMAATV VQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHCCFMAPYTHTQALAD SLRNMVQAEQENDTSKGFSREPMAMATV VQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHCCFMAPYTHTTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK: SLRNMVQAEQENDTSGWFDVLQKVSTQLK:	490	1840	A	3825	79	9748	
KSFQQQQQQQQQQQQQQQQQQQQQQQQQQPP PPPPPPPQIPPPPQIPPPPPPPP		1					
PPPPPPOL PO PPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDIA ESDVRMVADECLNKVIKALMDSNLFRLQLE. YKEIKKNGAFRSLRALWFAELAHLUSPQI CRPYLVNLLPCLTRTSKRPESSVGETLAAAV KIMASFGNFANDNEIKVLLKAFIANLKSSSPT RRTAAGSAVSICQHSRRTQYFYSWLLNVLLC LLVPVEDEHSTILLIGVLLTLRYLVPLLQQQ KDTSLKGSGGVTRKEWSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGIVEI AGGGSSCSPVLSRKQKGKVLLGEEBALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPROSHTULAQSVDLASCDLTSS ATDGDEDILSHSSSQVAVPSDPAAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE. FRNSSMALQQAHLLKNMSSICROPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTRTLTGNTFSL ACPULRRTLKDESSVTCKLACTAVRNCVM SLCSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFLYSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGBEDPRVRHVAAASI RLVPKLFYKCDQGQADPVWAVARDQSSVY KLLMHETQPSHFPSVSTITRLYGGVNLJPSITT VTMENNLSRVIAAVSHELTSTTRALTGCCC ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWPPLDLSAHQDAI ILAGNLLAASAPKSLRSWASSEEANPAATK QEEVWPALGDRALVPMVCQLFSHLLKVNIG AHVLDDVAPGPAIKAALPSLTRPPSLSPIRRK GKERPGEQASVPLSKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS GPVTTSKSSSLGSFYLSFKKGSEASAASRGSUS ATLODICKVEEILGTYLKSCFSREFMMATVC VQULKTLFGTNLASGFDGLSSNPSKSQGRA QRLGSSSVRGLYHYCFMAPYTHFTQALAD LSLRMMVQAEQEADDTSGWFDVLQKVSTQLKC		1	1	1		İ	ASRPEASGDCRAGRETAMATLEKLMKAFESL
GPAVAEËPLIRPKKELSATKKÖRVNHCLTUC ENIVAQSVRNSPEFQKLLGIAMELFILCSDDA ESSVRMVADECLNKVIKALMÖSNLPRLQLE YKEIKKINGAPRSLRAALWRFAELAHLVRPQI CRYLVVILLPCLTIKSKPESVVQETLAAAV KIMASFGNFANDNEIKVLLKAFIANLKSSSPT RRTAAGSAVSICQHSRRTQYFYSWLLNVLLC LLVPVEDEHSTLILLIGVLTLTRYLVPLLQQQ KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTIQHDHNVYTGALELLQQLFRTPPEI LQTLTAVGGIQQLTAAKEESGGRSRSGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVEISGELAASSGVSTP SAGHDITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQVTTEGFDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRISSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRKGDIGQSTDDDS APLVHCVRLLSASFLLTGGENVLVPPRADVS SVKALALSCVGAAVALHESSFSKLYKVPLL TTEVPEEQYVSDILNVDHGDPQVRGATAIL GTLICSILSRSPHIVGDWMGTRTLTGNTFSL ADCPPLIRKTLKDESSVTCKLACTAVRNCVA SLCSSSYSELGIQLIIDVLILRNSSYWLVRTE LETLAEDFRLVSFLKAEALHJRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI RLVPKLFYKCDQGQADPVVAARDQSSVY KLLMHETQPPSHSSVSTITRIYRGYNLLPSITL VTMENNLSRVIAAVSHELITSTTRALTGCCC ALCLLSTAPPCLWSFLSGWHGCVPPLSASDES KSCTVGMATMILTILLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASSEEANPAATK QEEVWPALGDRALVPNVCQLFSHLLKVNING AHVLDDVAPGPAIKAALPSLTNPSLSPIRK GKEEPGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHVDVLKATH NYKVTLDLQNNSTEKEGGFLRSALDVLSQILE ATLQDIGKCYEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRRGLYHYCFMAPYTHTTQALAD LSLRNMVQAEQEANDTSGWFDVLQKVSTQLK:		1	1		Ī		KSFQQQQQQQQQQQQQQQQQQQQQPPPP
ENIVAQSVRNSPEFQKLLGIAMELFILLCSDÜA ESDVRMVADECLNKVIKALMDSNLPRLQLE YKEIKKINGAPRSLRAALWRFAELAHLVRPQI CRPYLVNLLPCLTRTSKREESVQETLAAAV KIMASFGNFANDNEIVLKAFIANI.KSSSPT RRTAAGSAVSICOHSRRTQVFYSWLLNVLLC LLVPVEDEHSTLLILGVLLTLRYLVPLQQQ KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAKESGGGRSRSGSIVEL AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE FRNSSMALQQAHLLKNMSHCRQFSDSSVVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVYN SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYIDHGDPQVRGATAIL GTLICSILSRSRPHVGDWMGTRILTGNTFSL ACPULIKRTLKDESSVTCKLACTAVRNCVW SLCSSSYSELGLQLIIDVLTLRNSSYNLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILLGDEDPRVRHVAAASI RLVPKLFYKCDQGQADPVAVARDQSSVY KLIMHETQPSHPSSVSTITRIYRGYNLAPIT VTMENNLSRVIAAVSHELITSTTRALTFGCCC ALCLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLISSAWPPLDLSAHQDAI ILAGNLLAASAPKSLRSWASSEEANPAATK QEEWPALGDRALVPMVEQLFSHLLKVNIG AHVLDDVAPGPAIKAALPSLTRALFGCCI ALCLSTAFPVCIPSKERGESSAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCYEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQLALD LSLRNMVQAEQEANDTSGWFDVLQKSVTQLK:			İ	İ			PPPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPPPPPPP
ESDVRMVADECLNÉVIKALMDSNIPRLQLE YKEIKKNGAPRSLRAALWRFAELAHLVRPQI CRPYLVNLLPCLIRTSKRPESSVQETLAAAVI KIMASFGRFANDNEIKVLLKAFIANLKSSSTI RRTAAGSAVSICORSRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI THHTIQHQDHNVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLISHSSSQVSAYPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNYLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGQYVSDLNYIDHDDPQVRGATAILC GTLICSLSSRSPHVGDWMGTRTLTGHTFSL ADCIPLLRKTLKDESSVTCKLACTAYRNCVW SLCSSSYSSLGJQLUIDVLTRNSSYWLVRTE LETLAEIDFALVSSLEAKAENLHRGAHHYTG LKLQERVLNNVVHHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQQADPVVAVARDOSSVYV KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTIGCCI ALCLLSTAFPVCIWSLGWHGGVPLASPDIS KSCTVGMATMILTLSSAWPPLDLSAHDDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLESHLLKVINIC AHVLDDVAPGPAIKAALPSLTNIPPSLSPIRRK GKEKEPEGQASVPLSPKKGSEASAASRQSDI GPVTTSKSSSLGSFYHLPSYLKLHDUKATH NYKVTLDLQNSTEKFGGFIRSALDVLSQILE ATLQDIGKCVEEIGYLKSCFSREPFMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYHYCTMAPYTHETQALAD SLRNWVQAEGENDTSGWPDVLQKVSTQLK.		i	l				GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
ESDVRMVADECLNÉVIKALMDSNIPRLQLE YKEIKKNGAPRSLRAALWRFAELAHLVRPQI CRPYLVNLLPCLIRTSKRPESSVQETLAAAVI KIMASFGRFANDNEIKVLLKAFIANLKSSSTI RRTAAGSAVSICORSRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI THHTIQHQDHNVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLISHSSSQVSAYPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNYLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGQYVSDLNYIDHDDPQVRGATAILC GTLICSLSSRSPHVGDWMGTRTLTGHTFSL ADCIPLLRKTLKDESSVTCKLACTAYRNCVW SLCSSSYSSLGJQLUIDVLTRNSSYWLVRTE LETLAEIDFALVSSLEAKAENLHRGAHHYTG LKLQERVLNNVVHHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQQADPVVAVARDOSSVYV KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTIGCCI ALCLLSTAFPVCIWSLGWHGGVPLASPDIS KSCTVGMATMILTLSSAWPPLDLSAHDDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLESHLLKVINIC AHVLDDVAPGPAIKAALPSLTNIPPSLSPIRRK GKEKEPEGQASVPLSPKKGSEASAASRQSDI GPVTTSKSSSLGSFYHLPSYLKLHDUKATH NYKVTLDLQNSTEKFGGFIRSALDVLSQILE ATLQDIGKCVEEIGYLKSCFSREPFMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYHYCTMAPYTHETQALAD SLRNWVQAEGENDTSGWPDVLQKVSTQLK.		1		ľ	-		ENIVAOSVRNSPEFOKLLGIAMELFLLCSDDA
YKEIKKNGAPRSLRAALWRPAELAHLVRPQI CRPYLVNLIPCLTRTSKRPEESVQETLAAAVI KIMASFGMFANDNEIKVLLKAFIANILKSSST RRTAAGSAVSICQHSRRTGYFYSWLLNVLLG LLVPVPDEHSTLLIGVLTLRYLYLLQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVPSDDSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQAHLLKNMSHCRQPSDSSVDK VLRDEATEFGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGQYVSDIWYDHODPOVRGATALL GTLICSILSRSRPHVGDWMGTRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRIE LETLAEDFFLVSFLEAKABNLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI RLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVLAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHIGGVPPLSASDESS KSCTYGMATMILTILSSAWPPLDLSAHQDAI LAGNLAASAPKSLRSSWASEELBAPAATK QEEVWPALGDRALVPMVEQLFSHLKVINIC AHYLDDVAPGPAIKAALPSLTNIPPSLSPIRRK GKEKEPGGASVSKLRSSWASEEAANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHYLDDVAPGPAIKAALPSLTNIPPSLSPIRRK GKEKEPGGASVSKLRSSWASEEAANPAATK QEEVWFALGDRALVPMVEQLFSHLLKVINIC AHYLDDVAFGFAIKAALPSLTNIPPSLSPIRRK GKEKEPGGASVSSKLRSSWASEEAANPAATK QEEVWFALGDRALVPMVEQLFSHLLKVINIC AHYLDDVAFGFAIKAALPSLTNIPPSLSPIRRK GKEKEPGGGASVSKRSSMASEAASRQSDI GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQQGIGKCVEELIGYJKKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPRGLYYHCFMAPYTHFTQALAD SLRNWVQAEGENDTSGWFDVLQKVSTQLK.		1	)	ł		1	
CRPYL-VILLPCL.TRTSKPEESVQETLAAAV KIMASFGNFANDNEIKVLLKAFIANLKSSSPT RRTAAGSAVSICQHSRRTQYFYSWLLNVLLC LLVPVEDEHSTLLILGVLLTLRYL-VPLLQQQ KDTSLKGSFGVTKEMEVSPSAGDLQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKESGGRSRGSIVEI AGGGSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDLSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQVLGLQIGQPQEDEEATGLPDEASE, FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASLLTGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEQYVSDILNYIDHOPQVVRGATALLG GTLICSILSRSRFHVODWMGTRTLTGMTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCTW SLCSSSYSELGLQLIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHHLLGDEDPRVRHVAAASI RLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITL VTMENNLSRVIAAVSHELITSTTRALITGCCI ALCLLSTAFPVCIWSLGWHGGVPPLSASDES KSCTVGMATMILTLSSAWPPLDLSAHQDAI ILAGNLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHYLDDVAPGPAIKAALPSLTNIPPSLSPIRRK GKEKEPEGGASVPSKKGGEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE TALQDIGKCVEEIGYLKSCYSREPSPMATVC VQQLLKTLFGTNLASQPGGLSSNPSKSQGRA QRLGSSSVPGLYYHCFMAPTHHTQALAD SLRNMVQAEGENDTSGWFDVLQKVSTQLK.			1	ł	1		
KIMASFGNFANDNEIKVLIKAFIANIKSSSFT RRTAAGSAVSICQHSRRTQYFYSWLINVLIC LLVPVEDEHSTILUI.GVLITLRYLVPLIQQQ KDTSIKGSFGVTRKEMEVSPSAEQL VQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEI LQTLTAVGGIGQLTAAKEESGGRSRSGIVEI AGGGSSCSPVLSRKQKGKVLLGEEGALEDDS ESRSDVSSSA1 TASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDIL.SHSSSQVSAVPSDPAMDLNDG TQASSPISDSSOJTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE FRNSSMALQQAHLLKMSHCRQFSDSSVDK VLRDEATEPGDQENKPCRKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGYVYSDILNYDHDDPQVRGATAIL GTLICSILSRSRFHVGDWGTTRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVK SLCSSSYSELGQLIIDVLTLRNSSYWLVRTE LETLAEDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAAS IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLIMHETQPPSHFSVSTITRIVRGYNLPSITT VTMENNLSRVLAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLISAWFPLDLSAHQDA ILAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPRRK GKKEPGGGASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTILDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEGILGYYLKSCFSREPMMATVC VQQLLKTLFGTNLASOFDGLSSNPSKSQGRA QRLGSSSVRPGLYPYCFMAPYTHFTQALAD SLRNNWQAGENDTSGWFDVLQKVSTQLKS							
RRTAAGSAVSICQHSRRTQYFYSWLLNVLLC LLVPVEDEHSTILLIGVLLTLRYLVPLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEI LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRKGDIGGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGQYVSDILNYDHODPQVRGATAIL GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTILKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI RLVPKLFYKCDQGQADPVAVARDQSSVY KLLMHETQPPSHFSVSTITRYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAPPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTILLSSAWFPLDLSAHQDAJ ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVNIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKR GKEKEPGGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEELGYLKSCFSREPMMATVC VQQLLKTLFGTNLASGFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNNWQAGENDTSGWFDVLQKVSTQLKC							
LLVPVEDEHSTLLILGVLLT.RYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEI TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVEI AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDJITEQPRSQHTLQADSVDLASCDLTSS ATDGDEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGQDENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPSFPSKLYKVPLE TTEYPEEQYVSDILNYIDHODPQVRGATALL GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCPPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEDFRLVSFLEAKAENLHRGAHTHYTG LKILQERVLINNVVHHLLGDEPRVRHVAAAS IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAJ ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVING AHVLDDVARGPAKAALPSLTNPSLSFIRRK GKEKEPEGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQREQENDTSGWFDVLQKVSTQLKS					,	Į.	
KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEIT TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKESGGRSRSGIVEI AGGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGNVLVPDRDVRY. SVKALALSCVGAAVALHPESFFSKLYKVPLE TTEYPEGQYVSDILNYIDHGDPQVRGATALLC GTLICSILSRSRPHVGDWMGTRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVA. SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE. LETLAEDFRLVSPLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVFKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTTRYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCILSTAFPVCIWSLGWHCGVPPLSASDES. KSCTVGMATMILTLSSAWFPLDLSAHDDAI ILAGNLLAASAPKSLRSWASEEEANPAATK QEEVWPALGDRALYPMVEQLFSHLKVNIU. AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKC	ĺ			l	1		
TILHITQHONNVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGRSSGSUSTP AGGGSSCSPVLSRKQKGKVLLGEEFALEDDS ESRSDVSSALTASVKDEISGELAASSGVSTP SAGHDIITQPRSOHTLQADSVDLASCDLTSS ATDGDEEDI.SHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEFATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRI SVKALALSCVGAAVALHPESFFSKLYKVPL TTEYPEEQYVSDILNYIDHGDPQVRGATAIL. GTLICSILSRSRPHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVW SLCSSSYSELGQLIIDVLTRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHILGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIVRGYNLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPPWCQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLAQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLKC	ļ		1	j	]		· · · · · · · · · · · · · · · · · · ·
LQTLTAVGGIGQLTAAKEESGGRSRSGIVEI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEQPRSQHTI.QADSVDLASCDLTSS ATDGDEEDILSHSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSMALQQAHLLKNMSHCQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRN SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYDDHGDPPQVRGATALL GTLICSILSRSRPHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRTYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI LAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK		1	i				
AGGSSCSPVLSRKQKGKVLLGEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE, FRNSSMALQQAHILKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYIDHGDPQVRGATALL GTILCSLSRSRPHVGDWMGTIRTLTGNTTSL ADCIPLLRKTLKDESSYTCKLACTAVRNCVW SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVLAAVSHELITSTTRALIFGCCI ALCILSTAFPVCIWSLGWHGGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHYLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK		ı		!			TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEGPRSQHTILQADSVDLASCDLTSS ATDGDEEDII.SHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYIDHGDPQVRGATAILG. GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTE. LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRLYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPGGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWPDVLQKVSTQLKC				i			LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
ESRSDVSSSALTASVKDEISGELAASSGVSTP SAGHDIITEGPRSQHTILQADSVDLASCDLTSS ATDGDEEDII.SHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYIDHGDPQVRGATAILG. GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGQLIIDVLTLRNSSYWLVRTE. LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRLYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPGGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWPDVLQKVSTQLKC							AGGGSSCSPVLSRKOKGKVLLGEEEALEDDS
SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQFSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLE TTEYPEQYVSDILNYIDHGDPQVRGATALLC GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LEITLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHILLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALIFGCC ALCLLSTAFFVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHILKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLKC	ľ	1	1	l			·
ATDGDEEDĪLSHSSQVŠAVPSDPAMDLNDG TQASSPISDSQTITTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLD TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTRILTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVW SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVFKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAJ ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSFKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLKT		1	1	[		1	
TQASSPISDSSQTTTEĞPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEATGILPDEASE. FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFFGGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK		ł	1	1		}	
GTDNQYLGLQIGQPQDEDEEATGILPDEASE. FRNSSMALQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVXI SVKALALSCVGAAVALHPESFFSKLYKVPLE TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVW SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFFGGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK		i	1				
FRNSSMALQQAHLLKNMSHCRQPSDSSVDK VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLL TTTEYPEGQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTRALTFGCCE ALCILSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTILSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKPPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK		1				1	
VLRDEATEPĞDQENKPCRIKGDİĞQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRI SVKALALSCVGAAVALHPESFFSKLYKVPLL TTEYPEGYVSDİLNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIQ AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYK VTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHVCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK				1	1		
APLVHCVRLLSASFLLTGGKNVLVPDRDVRY SVKALALSCVGAAVALHPESFFSKLYKVPLE TTEYPEGYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRILTGNTFSL ADCIPLIRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALITGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES. KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHVCFMAPYTHFTQALAD SLRNMVQAEQENDTSGWFDVLQKVSTQLK	}	1	}	1	}		
SVKALALSCVGAAVALHPESFFSKLYKVPLE TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADL SLRNMVQAEQENDTSGWFDVLQKVSTQLK					1		VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES. KSCTVGMATMILTLLSSAWFPLDLSAHQDAJ ILAGNLLAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK				1	1		
TTEYPEEQYVSDILNYIDHGDPQVRGATAILG GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCLLSTAFPVCIWSLGWHCGVPPLSASDES. KSCTVGMATMILTLLSSAWFPLDLSAHQDAJ ILAGNLLAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSVLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK	1		l	1	1		SVKALALSCVGAAVALHPESFFSKLYKVPLD
GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAAS) IRLVPKLFYKCDQGQADPVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCI ALCILSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK				1	1	1	TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
ADCIPLRKTLKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALIFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTILLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1		1	1	1		GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTE LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALIFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	ļ				1		*
LETLAEIDFRLVSFLEAKAENLHRGAHHYTG LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITI VTMENNLSRVIAAVSHELITSTTRALIFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIO AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVO VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	}		1		1	<u> </u>	* - *
LKLQERVLNNVVIHLLGDEDPRVRHVAAASI IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALIFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTILLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1	1	1	1		
IRLVPKLFYKCDQGQADPVVAVARDQSSVY KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	ļ	ļ	Į.	Į	1		
KLLMHETQPPSHFSVSTITRIYRGYNLLPSITT VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT					1		
VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTLLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT			1	Į	1		
ALCLLSTAFPVCIWSLGWHCGVPPLSASDES KSCTVGMATMILTILLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1						
KSCTVGMATMILTILLSSAWFPLDLSAHQDAI ILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT					1		VTMENNLSRVIAAVSHELITSTTRALTFGCCE
KSCTVGMATMILTILLSSAWFPLDLSAHQDAI IILAGNLLAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	1	1				ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
ILAGNLLAASAPKSLRSSWASEEEANPAÄTK QEEVWPALGDRALVPMVEQLFSHLLKVINIG AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT	1	}	1	J	j	]	KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK			1		1		1 ·
AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK	[		1	1	1		
GKEKEPGEQASVPLSPKKGSEASAASRQSDT GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK	1	}					
GPVTTSKSSSLGSFYHLPSYLKLHDVLKATH NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK			1		1		
NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK	]	1	İ	1	1		
ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK			1		l		
VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLKT		Į.				}	
QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK	1	1		1	I '		ATLQDIGKCVEEILGYLKSCFSREPMMATVC
QRLGSSSVRPGLYHYCFMAPYTHFTQALAD. SLRNMVQAEQENDTSGWFDVLQKVSTQLK		1				Ì	
SLRNMVQAEQENDTSGWFDVLQKVSTQLK			1				
	1		1		}		
I I I I I I I I I I I I I I I I I I I	1	i		1	1		
AND AND TO SELECT THE CONTROL OF THE	L	<u> </u>			L	L	TATTO A TEST TOTAL

SEO ID   SEQ ID   Met   SEO   Predicted   Predicted end   Amino acid sequence (A=Alanine C=Cysteine.	SEQ ID   SEQ ID   Met   SEQ   Predicted   Predicted end   Amino acid sequence (A=Alanine C=Cysteine,
NO. of mucl-cotide sequence    No. of mucl-cotide sequence	NO: of noted and note of the noted to the noted to the noted of the no
NO. of mucl-cotide sequence    No. of mucl-cotide sequence	No. of   No. of   No. of   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide   coation   nucleotide
NO. of mucl-cotide sequence where the carried coation in cuclotide processing in the carried coation of the carried sequence where the carried sequence where the carried sequence sequence where the carried sequence sequence where the carried sequence sequence where the carried sequence sequence where the carried sequence sequ	NO: of No: of No: of nucleotide colide sequence when the colide sequence of the colide of the colide of the colide of the colide of the colide of the colide of the colide of the colide of the colide of the colide of the colider of peptide sequence of the colider of peptide sequence of the colider of the c
NO: of mucl- coide out of the control of the contro	NO: of No: of No: of nucleotide output of the control of the contr
NO. of No. po de noucle could could noucle could sequence wence where the country of the country	NO: of mucicolide colide sequence USSN plants Acid, E-Glutamic Acid, USSN plants Acid, Sequence USSN plants Acid, Sequence USSN plants Acid, Sequence USSN plants Acid, Sequence USSN plants Acid, Sequence USSN plants Acid, Sequence peptide sequence peptide residue of peptide peptide sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition Landson, Sequence Proposition, Seque
NO: of Mo: of mucl-coide coide uence where the control of the coide of	NOC of Not of nucleotide colde sequence    D NO: per period point of the period color of the period sequence   D-Asyartic Acid, E-Glutamic Acid,   D-Asyartic Acid, E-Glutamic Acid,   D-Asyartic Acid,   E-Glutamic Acid,   D-Asyartic Acid,   E-Glutamic Acid,   D-Asyartic Acid,   E-Glutamic Acid,   D-Asyartic Acid,   E-Glutamic Acid,   D-Asyartic Acid,   E-Glutamic,   D-Asyartic Acid,   E-Glutamic,   D-Asyartic,   D-Asyar
NO: of No: of mucleotide outside of the property of the proper	NOC of Not of nucleotide colde order in the cold of nucleotide colded order of the cold of
NO: of No: of nucle colide sq- mucle colide sq- uence unce unce unce unce unce unce unce	NOC of Not of nucleotide order of the control of th
NO: of No: of nucle colide sq- mucle colide sq- uence unce unce unce unce unce unce unce	NO: of nucleotide order of the property of the control of the property of the
NO: of No: of No: of neucleotide colde sequence when the neucleotide sequence uence	NO: of nucleotide ocation under the control of period ocation of period ocation of period ocation of period ocation of period ocation of period ocation of period ocation of period ocation ocation of period ocation
NO: of No: of nucleotide olide sequence    No: of public peptide olide sequence   USSN   09/496   1914	NO: of nucle order of the property of the prop
NO: of No: of nucle ecide sequence    Daylor   D	NO: of NO: of nucleotide pulse ocitide sequence where the nucleotide cottide sequence where the nucleotide plant is not sequence where the nucleotide sequence where the nucleotide sequence where the nucleotide sequence where the nucleotide sequence where the nucleotide deletion, "Passarigine, P-Proline, M-Meditionine, N-Asparagine, P-Proline, M-Proline, N-Proline, N-Proline, N-Proline, N-Proline, N-Prol
NO: of No: of nucle coide sequence USSN 99/496 (USSN 99/496) uence	NO: of peptide sequence wence where the peptide sequence wen
NO: of nucle ordication   Di No: of in unucleotide   Despite   No: of in USNN   O9/496	NO: of No: of nucle cotide sequence where the contract of the contract of the cotide sequence where the cotide sequence wh
NO: of nucle colide sequence   Di No: of in uncleotide peptide sequence   S	NO of peptide cotide sequence    NO of peptide cotide sequence
NO: of nucle ocide sequence peptide could sequence under the content of the conte	NO: of nucle ordice colide sequence    No: of peptide sequence
NO: of nucle ocide sequence peptide could sequence under the content of the conte	NO: of nucle ordice colide sequence    No: of peptide sequence
NO: of nucle colide sequence   Di No: of in uncleotide nucleotide sequence   Sequence	NO of peptide cotide sequence    No of peptide cotide sequence
NO: of peptide cotide sequence    NO: of peptide cotide sequence	NO: of NO: of peptide cotide sequence    No: of peptide sequence
NO: of peptide cotide sequence    No: of peptide cotide sequence   Sequence	NO of nucle peptide sequence uence
NO: of peptide cotide sequence    No: of peptide cotide sequence	NO of nucle peptide cotide sequence uence
NO: of NO: of peptide cotide octide cotide sequence u	NO of nucleotide peptide cotide sequence uence used provide sequence used provides and provide sequence used provides and provi
NO: of NO: of peptide cotide cotide sequence ue	NO of nucle otide peptide cotide sequence uence USSN location 09/496 or peptide residue of peptide residue of peptide sequence se
NO of NO of peptide cotide cotide cotide cotide cotide cotide sequence u	NO of nucleotide peptide sequence uence use of peptide sequence uence uence use of peptide sequence uence use of peptide sequence use of peptide sequence use of peptide residue of peptide residue of peptide sequence sequence use of peptide residue of peptide sequenc
NO: of NO: of nucleotide peptide cotide sequence uence	NO: of nucleotide peptide cotide sequence  NO: of peptide cotide sequence  NO: of nucleotide peptide sequence  NO: of peptide sequence  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  NO: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (orresponding to last amino acid residue of peptide sequence)  No: of nucleotide location (nucleot
NO of NO of peptide cotide cotide cotide cotide cotide cotide sequence u	NO: of nucle peptide cotide sequence
NO: of nucleotide peptide sequence uence USN 09/496 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 915 greater amino acid residue of peptide sequence 915 greater amino acid residue of peptide sequence 916 greater amino acid residue of peptide sequence 917 greater amino acid residue of peptide sequence 917 greater amino acid residue of peptide sequence 918 g	NO: of nucle peptide cotide sequence
NO: of nucleotide peptide sequence uence USN 09/496 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 914 greater amino acid residue of peptide sequence 915 greater amino acid residue of peptide sequence 915 greater amino acid residue of peptide sequence 916 greater amino acid residue of peptide sequence 917 greater amino acid residue of peptide sequence 917 greater amino acid residue of peptide sequence 918 g	NO: of nucle peptide cotide sequence
NO: of nucleotide peptide sequence	NO: of nucleotide cotide sequence peptide sequence of nucleotide location of peptide sequence of seque
NO: of nucleotide peptide sequence uence	NO: of nucleotide cotide sequence peptide sequence uence  NO: of nucleotide peptide sequence sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of the peptide sequence uence uence  NO: of nucleotide location of the peptide sequence uence uence uence uence  No: of nucleotide location of the peptide sequence uence
NO: of nucleotide cotide sequence peptide sequence uen	NO: of nucleotide cotide sequence peptide sequence peptide sequence peptide sequence punce unce nucleotide location sequence peptide sequence
NO: of nucleotide peptide sequence peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence uence  NO: of peptide sequence sequence  NO: of nucleotide location of peptide sequence uence  NO: of nucleotide location of peptide sequence sequence sequence  NO: of nucleotide sequence sequence sequence sequence sequence sequence  NO: of nucleotide sequence seque	NO: of nucl- peptide cotide sequence  NO: of nucl- peptide cotide sequence  NO: of nucleotide location location or seponding sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location location or sequence  NO: of nucleotide location or sequence  No: of peptide sequence  No: of nucleotide location location or sequence  No: of peptide sequence  No: of p
NO: of nucle peptide cotide sequence	NO: of nucl- peptide cotide cotide cotide cotide seq- uence  NO: of nucl- peptide cotide cotide seq- uence  NO: of nucl- peptide cotide seq- uence  NO: of nucl- peptide cotide seq- uence  NO: of nucleotide location location corresponding to last amino acid residue of peptide residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  No: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence of last amino acid residue of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of
NO: of nucle peptide cotide sequence	NO: of nucl- peptide cotide cotide cotide cotide seq- uence  NO: of nucl- peptide cotide cotide seq- uence  NO: of nucl- peptide cotide seq- uence  NO: of nucl- peptide cotide seq- uence  NO: of nucleotide location location corresponding to last amino acid residue of peptide residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  No: of nucleotide incurrence in location corresponding to last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence in last amino acid residue of peptide sequence  No: of nucleotide incurrence of last amino acid residue of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of peptide sequence  No: of nucleotide incurrence of
NO: of nucle peptide cotide sequence  NO: of nucle peptide sed cotide sequence  NO: of nucleotide location location of sequence  NO: of nucleotide sequence  No: of sequence  NO: of nucleotide sequence  No: of sequence  NO: of nucleotide sequence  No: of sequence  No: of sequence  NO: of nucleotide sequence  No: of seq	NO: of nucle peptide cotide sequence  NO: of nucleotide peptide cotide sequence  NO: of nucleotide peptide cotide sequence  NO: of nucleotide peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location corresponding to last amino acid residue of peptide sequence  No: of nucleotide location (P-possible nucleotide deletion, N-possible nucleotide locatio
NO: of nucleotide cotide sequence  NO: of nucleotide peptide cotide sequence  NO: of nucleotide location in uncleotide location sequence  NO: of nucleotide peptide sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location sequence  NO: of nucleotide location in location in location in location in location acid residue a maino acid residue of peptide residue of peptide sequence  NO: of nucleotide location in locat	NO: of nucle cotide cotide sequence wence
NO: of nucle- vertice cotide sequence  NO: of nucleotide sequence  NO: of peptide cotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide s	NO: of nucl- peptide cotide sequence uence of peptide sequence outled sequence of peptide of peptide
NO: of nucle- vertice cotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide location orresponding to last amino acid residue of peptide sequence  NO: of nucleotide location  NO: of nucleotide location  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of nucleotide location  NO: of nucleotide location  NO: of peptide sequence  NO: of nucleotide location  NO: of nucleotide location  NO: of peptide sequence  NO: of nucleotide location  No: of nucleot	NO: of nucl- peptide cotide sequence uence of peptide sequence outled sequence of peptide of peptide
NO: of nucl- original in uncleotide sequence  NO: of nucleotid- sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide location uncleotide location orresponding to last amino acid residue of peptide sequence  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide sequence  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  NO: of nucleotide location  No: of nucleotide l	NO: of nucl- otide sequence wence, wence, we wence, we want with a warring wence, we will be a warring with a warring wence, we want was a warring warring was a warring w
NO: of nucl- otide sequence wence, wenc	NO: of nucl- cotide sequence w
NO: of nucle of nucle cotide sequence  NO: of peptide sequence  No: of peptide sequence  No: of	NO: of nucleotide cotide sequence  NO: of peptide cotide sequence  NO: of peptide cotide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide
NO: of nucl- cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide s	NO: of nucleotide cotide sequence  NO: of nucleotide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide sequence  N
NO: of nucl- cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide s	NO: of nucleotide cotide sequence  NO: of nucleotide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide sequence  N
NO: of nucleotide cotide sequence  NO: of peptide cotide sequence  NO: of peptide cotide sequence  NO: of peptide cotide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide to cation on g to first amino acid residue of peptide sequence  No: of peptide sequence  NO: of peptide sequence  No: of peptide to cation on g to first amino acid residue of peptide sequence  No: of peptide to corresponding to last amino acid residue of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide to cation ocorresponding to last amino acid residue of peptide sequence  No: of peptide seq	NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide sequen
NO: of nucl- cotide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- No: of peptide sequence  No: of peptide sequenc	NO: of nucleotide cotide sequence uence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide sequ
NO: of nucl- ectide seq- uence  NO: of nucl- ectide seq- uence  NO: of peptide seq- uence  NO: of nucl- ectide seq- uence  NO: of nucl- ectide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide sequence  No: of peptide seq- No: of peptide sequence  No: of peptide sequence  No: of peptide seq- No: of peptide sequence  No: of peptide se	NO: of nucle- cotide seq- uence uence uence    NO: of nucleotide cotide seq- uence uence    NO: of nucleotide cotide seq- uence uence    NO: of nucleotide cotide seq- uence uence    NO: of nucleotide seq- uence    NO: of peptide seq- uence    NO: of nucleotide seq- uence    NO: of peptide seq- uence    No: of peptide seq- uence    No: of peptide seq- uence    No: of peptide seq- uence    No: of peptide seq- uence    No: of peptide seq- uence    No: of peptide seq- uence    No: of peptide sequence    No: of pep
NO: of nucleotide peptide eotide sequence uence	NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide sequence  No: of peptide sequence  No: of peptide sequence  Og/Ap6  Sequence  No: of peptide sequence  Og/Ap6  Og/Apf  Og/Apf  Og/Apf  Og/Apf  Og/Apf  Og/Apf  Og/Apf  Og/Apf  O
NO: of nucleotide peptide eotide sequence uence	NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide cotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  No: of peptide sequence  No: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  OS: of peptide sequence  N=Methionine, N=Asparagine, P=Proline, Q=Glutamine, C=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, C=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tryrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \= possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLVQLVLLLUDSQLVQLVLLLUDSQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQ
NO: of nucleotide eotide sequence  NO: of nucleotide eotide sequence  NO: of nucleotide eotide sequence  NO: of nucleotide eotide sequence  NO: of nucleotide eotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NE-MENDAMENTAL REPROBLEM SEQUENCE  NE-Phenylalanine, G=Glycine, H=Histidine, lessolucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHDSE	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  No: of peptide sequence  NE-Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, M=Isoleuche, M=Isoleuche, N=Isoleuche, M=Isoleuche, N=Isoleuche, M=Isoleuche, N=Isoleuche, N=Isoleuche, N=Isoleuche, N=Isoleuche,
NO: of nuclectide cetide sequence  NO: of nuclectide cetide sequence  NO: of nuclectide cetide sequence  NO: of nuclectide cetide sequence  NO: of nuclectide cetide sequence  NO: of nuclectide sequence  NO: of nuclectide sequence  NO: of peptide sequence  NO: of nuclectide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of pepti	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  No: of peptide sequence  No: M=Methionine, N=Asparagine, P=Proline, Q=Glutamice, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, Y=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFIFLYLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMIII DSHEALGVLNTLFEILAPSSLRPVDMLLRSM VTPNTMASVSTVQL WISGILAILRVLISQSTEI IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFK GMFRITAAATRLFRSDGCGGSFYTLDSLNL ARSMITTHPALVLLWCQILLLVNHTTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDL
NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NE-Methionine, N=Asparagine, P=Proline, M=Methionine, N=Asparagine, P=Proline, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Methioni	NO: of nucleotide cotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence  No: of nucleotide sequence, sequence  No: of nucleotide sequence, sequence, sequence  No: of nucleotide sequence, sequence, sequence, sequence  No: of nucleotide sequence, sequence, sequence, sequence, sequence  No: of nucleotide sequence, sequence, sequence, sequence, sequence  No: of nucleotide sequence, sequence, sequence, sequence, sequence, sequence  No: of nucleotide sequence, se
NO: of nucleotide cotide sequence  NO: of nucleotide sequence  No: of peptide sequence  N=Methionine, N=Asparagine, P=Proline, W=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=roline, N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine,	NO: of nucl- ocide seq- uence  NO: of peptide eotide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  No: of peptide sequence  No: of peptide s
NO: of nucleotide cotide sequence  NO: of nucleotide sequence  No: of peptide sequence  N=Methionine, N=Asparagine, P=Proline, W=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=Proline, N=Methionine, N=Asparagine, P=roline, N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Stop N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine, N=Methionine,	NO: of nucl- ocide seq- uence  NO: of peptide eotide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  No: of peptide sequence  No: of peptide s
NO: of nucleotide sequence  No: of nucleotide sequence  N=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, S=Serine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, S=Serine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, S=Serine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Serine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, R=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, R=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Histidine, S=Isoleucine, R=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Methionine, N=Asparagine, P=Proline, Q=Glycine, H=Methionine, N=Isoleucine, Sequence Sequence Sequence Sequence Sequence Sequence Sequence Seq	NO: of nucleotide cotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of nucleotide sequence  No: of peptide sequence
NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide seq	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  No: of pepti
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- No: of peptide seq- No: of peptide seq- No: of peptide seq- No: of peptide seq- No: of peptide seq- No: of peptide seq- No: of p	NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  N=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFIFILAYENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLRSMIVENCENEDKWKRLSRQIADIILPMLAKQQMFILARVLISQSTEILFIVENCENEDKWKRLSRQIADIILPMLAKQQMFILARVLISQSTEILFIVENCENEDKWKRLSRQIADIILPMLAKQQMFILARVLISQSTEILFIVENCENEDKWKRLSRQIADIILPMLAKQQMFILARVLISQSTEILFIVENCENEDKWKRLSRQIADIILPMLAKQUMRIVARVLISQSTEILFIVENCENEDKWKRLSRQIADIILPMLAKQUMR
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptid	NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  N=Phenylalanine, G=Glycine, H=Histidine, lelsoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFIFILATELY SEQUENCE  N=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFIFILATELY SEQUENCE  N=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLVQLVLDLLAQLVQLVQLVQLVLDLLAQLVQLVQLVQLVQLVLDLLAQLVQLVQLVQLVQLVQLVQLVQLVLDLLAQLVQLVQLVQLVLDLLAQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQLVQ
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptid	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  NeMethionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFIFILAYSTERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMI
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptid	NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide seq
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptid	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of pepti
NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptid	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of pepti
NO: of nucleotide peptide eotide sequence  NO: of nucleotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide	NO: of nucl- octide seq- uence  NO: of peptide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of nucl- octide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uence  N=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFI FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
NO: of nucl- operatide seq- uence  NO: of peptide seq- uence  NO: of nucl- operatide seq- uence  NO: of peptide seq- uence  NO: of nucl- operatide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uenc	NO: of nucl- ocide seq- uence  NO: of peptide seq- uence  NO: of nucl- ocide seq- uence  NO: of nucl- ocide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of pe
NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NE-Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF	NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of pe
NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide sequence  NO: of nucleotide peptide sequence  NO: of nucleotide peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NE-Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion	NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of nucl- otide seq- uence  NO: of nucl- otide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  No: of peptide seq- uence  Ne-Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of nucleotide peptide eotide sequence  NO: of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NEMENDENTIAL RESIDIATION (SECTION NEMENTAL SECTION NEMEDIAL SECTION NEM	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  Ne=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
NO: of nucleotide cotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NO: of peptide in nucleotide location corresponding to last amino acid residue of peptide sequence  NEMENTATIONAL DESIGNATION SESSION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATIONAL DESIGNATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)  NEMENTATION (NEMENTAL PROSIBLE)	NO: of nucleotide sequence  NO: of peptide sequence  NO: of nucleotide sequence  NO: of nucleotide sequence  NO: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  No: of peptide sequence  Ne Methionine, Ne Asparagine, Pe Proline, Qe Glutamine, Re Arginine, Se Serine, Te Threonine, Ve Valine, We Tryptophan, Ye Tyrosine, X = Unknown, * = Stop codon,
NO: of nucl- eotide seq- uence    NO: of nucl- eotide seq- uence   NO: of nucleotide	NO: of nucl- eotide seq- uence  NO: of peptide seq- uence  NO: of nucl- eotide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  NO: of peptide seq- uence  No: of pept
NO: of nucl- eotide seq- uence USSN location correspondi ng to first amino acid residue location graphic amino acid residue location corresponding to last amino acid residue location corresponding	NO: of nucl- cotide seq- uence  NO: of peptide seq- uence  NO: of poptide seq- uence  No: of poptide s
NO: of nucl- eotide seq- uence USSN location correspondi to last amino uence USSN location gine first location golden location correspondi to last amino acid residue location golden location corresponding to last amino acid residue location corresponding last amino acid residue location corresponding to last amino acid residue location corresponding last amino acid residue location corresponding location location corresponding location location location location location corresponding location	NO: of nucl- cotide seq- uence uence uence NO: of nucle of the period in nucleotide seq- uence uence nucleotide location corresponding to last amino ng to first nucleotide location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino location corresponding to last amino acid residue location corresponding to last amino
NO: of nucl- eotide seq- uence USSN location correspondi to last amino uence USSN location gine first location golden location correspondi to last amino acid residue location golden location corresponding to last amino acid residue location corresponding last amino acid residue location corresponding to last amino acid residue location corresponding last amino acid residue location corresponding location location corresponding location location location location location corresponding location	NO: of nucl- cotide seq- uence uence uence NO: of nucle of the period in nucleotide seq- uence uence nucleotide location corresponding to last amino ng to first nucleotide location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino acid residue location corresponding to last amino location corresponding to last amino acid residue location corresponding to last amino
NO: of nucl- eotide seq- uence	NO: of nucl- peptide eotide seq- uence O9/496 correspondi to last amino NO: of nucle oide seq- uence NO: of nucle oide seq- uence NO: of nucle oide seq- uence NO: of nucleotide location corresponding to last amino NO: of peptide in nucleotide location corresponding to last amino NEMethionine, N=Asparagine, P=Proline,
NO: of NO: of peptide in nucleotide location per la location per la location per la location location location per la location per la location per la location location per la location per la location location location per la location location la	NO: of NO: of peptide on the nucleotide of the n
NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glutamic Acid, nucleotide location F=Phenylalanine, G=Glycine, H=Histidine,	NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glutamic Acid, nucleotide location F=Phenylalanine, G=Glycine, H=Histidine,
NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glutamic Acid,	NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glutamic Acid,
	The contract of the contract o

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
493	1045		3636		300	KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG FTVLARMVLIS*PCDPPTLASQGTAITGMSYH ARPQDIDFLYAHQGRCWFRLL
495	1845		3847	1774	40	DIFFRAKEGMGQDEAQFSVEMPLTGKAYL WADKYRPKPRFFNRVHTGFEWNKYNQTHY DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL EACADNKDFAILRFHAGPPYEDIAFKIVNREW EYSHRHGFRCQFANGIFQLWFHFKRYRYRR* RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC HGELRRHWDRLA*GPDATEGALGASFEHEG GQQPPADLTVQADTLHRPSARLGGAHRACPK RRPHRVLWRWARGAWAWRCQAREKQETQG QPCHITGHPLGREAEPAAAGAAPALAHRPPF ARTGSTEVPGPCWRPIRHCRRDPLWTPTLCND WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN VMGTKSH*AVLPPPPSTGPGQGLPEGWGLE KGEGLPPGIPPGLLTGPWSMRPVTPSFAHIR TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK SFVLMELAYWQDRMFF
496	1846	A	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG LKLLTSSALPALASQSAEITGMSHRIWPLPLLR RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	A	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL
498	1848	A	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF\*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	A	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

	T 020		000			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		] .		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	<del>                                     </del>	<del>                                     </del>				APDPKALRLGVLKKKAMLHQEGHMDDALSL
		!				TRCQQEESQAARMIHSTNGLYNOFIKSLDSFS
	1	1				GKPRGSGPPAGTALPIEGVILSLODLIIYFEPPS
	1					EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV
		1	,			LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
		1				VNLLYELLASLIRGNRSNCALFSTNLDWLVS
				ĺ		l
	1					KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
						RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
	1					IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
†		1				ATHLRVGWALTEGYTPYPGAGEGWGGNGV
	1	ļ				GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
1	1 .	1		}	1	LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
						NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
	Į.	l		· ·		LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
	1					RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
	1				l	LERIREKLAENIHELWALTRIEQGWTYGPVRD
1	ł	ł				DNKRLHPCLVDFHSLPEPERNYNLQMSGETL
1						KTLLALGCHVGMADEKAEDNLKKTKLPKTY
ł	ł					MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
1						NGHNVWARDRVGQGWSYSAVQDIPARRNPR
i		ì				LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
						QSGRWYFEFEAVTTGEMRVGWARPELRPDV
						ELGADELAYVFNGHRGQRWHLGSEPFGRPW
		1		,		QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
		1				ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
		}				VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
1						KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
1		[			1	LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
		1				CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
1		}				NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
		1				AGGEAQPARAENEKDATTEKNKKRGFLFKA
						KKVAMMTQPPATPTLPRLPHDVVPADNRDD
						PEILNTTTYYYSVRVFAGQEPSCVWAGWVT
1	!				1	PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
		}				HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
		1				VIGCLVDLATGLMTFTANGKESNTFFQVEPN
		1				TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
1		]				AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
}		1			ł	MPNHFLQVETRRAGERLGWAVQCQEPLTMM
		İ				ALHIPEENRCMDILELSERLDLQRFHSHTLRL
						YRAVCALGNNRVAHALCSHVDQAQLLHALE
						DAHLPGPLRAGYYDLLISIHLESACRSRRSML
1		1	1		ĺ	SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
i		1				GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
		ł	1			ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
1		I			]	RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
1		1	1			DVKQILKMIEPEVFTEEEEEEDEEEGEEEDEE
		l			•	EKEEDEETAQEKEDEEKEEEAAEGEKEEG
		1				LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
		1		1	l	HRVESLAAFAERYVDKLQANQRSRYGLLIKA
						FSMTAAETARRTREFRSPPQEQINMLLQFKDG
1		1				TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
1						GEEEPEEETTLGSRLMSLLEKVRLVKKKEEK
1	1	l	İ			PEEERSAEESKPRSLQELVSHMVVRWAQEDF
						VQSPELVRAMFSLLHRQYDGLGELLRALPRA
						YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
1				ļ		ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
1	1	I	l	1		TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL
	1		3			

CEOTO	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID			•	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:			F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in	nucleotide	location	
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	j		ł	peptide	•	/=possible nucleotide deletion, \=possible
			i	sequence		nucleotide insertion
		<u> </u>	<del></del>	Sequence	<del></del>	CYFCRISRQNQRSMFDHLSYLLENSGIGLGM
1						QGSTPLDVAAASVIDNNELALALQEQDLEKV
	1					VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
1	ł		1			ERYLDFLRFAVFVNGESVEENANVVVRLLIR
						KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
			ļ.			DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
1			ł			FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
			l			I DOLADI EDI ACHOI DI CIDEI CADCAL ACADA
						LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
Į.	ì	Ì	ļ			MSASFVPDHKASMVLFLDRVYGIENQDFLLH
1		l				VLDVGFLPDMRAAASLDTATFSTTEMALAV
1						NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
	!	ł	1			MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
1	1	ŀ	1	l		YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
1		1				TNHYERCWKYYCLPTGWANFGVTSEEELHL
[	1	[	}			l
1	l .	1	1			TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
1	· .	ĺ	1			AIAGALPPDYVDASYSSKAEKKATVDAEGNF
		l	l			DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
			1		}	DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
	l	l	i			KDKETYRWPIKESLKAMIAWEWTIEKAREGE
1	1	1	i			EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
ł	l	ł	ł		ł	SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
Ī		l	ł		ļ	
1		l			1	ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
1	l					QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
		i				FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1	ļ			}	ĺ	EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
f		ĺ	ĺ	1	ĺ	TPAKVLGSGGHASNKEKEMITSLFCKLAALV
						RHRVSLFGTDAPAVVNCLHILARSLDARTVM
			1		١ .	KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
-						
ļ						KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
1	j.	1	į.	}		QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
	ł				1	TTKNTYVEKLRPALGECLARLAAAMPVAFLE
					Ì	PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
		ì				CPDIPVLERLMADIGGLAESGARYTEMPHVIE
					•	ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
						CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
1	1	l	l	Į.	1	KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
1	]	l	1			
	1	!	1			AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
1	1	İ	l			VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
1		ļ		]	<b>}</b>	AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
		ļ	1		1	NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
		1	1	i	1	RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
J		J	ļ	J	Į.	MCAPTDQDLITLAKTRYALKDTDEEVREFLH
1					1	NNLHLQGKVEGSPSLRWQMALYRGVPGREE
1	l .			1	1	DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
1	1	1	1		}	
1		l			I	KKAVWHKLLSKQRRRAVVACFRMTPLYNLP
1	1	1			ļ	THRACNMFLESYKAAWILTEDHSFEDRMIDD
1		I	<b>!</b>	ł	1	LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
					İ	ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
				}	I	GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
1				ł	1	HTRGAAEMVLOMISACKGETGAMVSSTLKL
				1	1	
1	'			1	[	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
1	1	1	1	i	1	IQALMQTCSVLDLNAFERQNKAEGLGMVNE
			1		1	DGTVINRQNGEKVMADDEFTQDLFRFLQLLC
1	1	1	1	1		EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL
Į.		1				RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
1	}	1	[	1	· ·	SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
Į.		j	J	}	1	DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
[				1		LDLQKDMVVMLLSLLEGNVVNGMIARQMV
1				1		
				1		DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
	1		<u> </u>	<u> </u>	L	QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI
						· · · · · · · · · · · · · · · · · · ·

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
		Ì		sequence		nucleotide insertion
		<u> </u>				QFLLSCSEADENEMINCEEFANRFQEPARDIG
		ŀ	1			FNVAVLLTNLSEHVPHDPRLHNFLELAESILE
		ŀ	ĺ			YFRPYLGRIEIMGASRRIERIYFEISETNRAQW
				i	1	EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
	į					FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA
	}		1			AEAGAEGAEEGAAGLEGTAATAAAGATARV
		l	ŀ			VAAAGRALRGLSYRSLRRRVRRLRRLTAREA
		1	ļ			ATAVAALLWAAVTRAGAAGAGAAAGALGL
	1	ľ				LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
				1		SDEVHGEQPAGPGGDADGEGASEGAGDAAE
						GAGDEEEAVHEAGPGGADGAVAVTDGGPFR
		i				PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG
					l	VDGVEEELPPEPEPEPEPELEPEKADAENGEK
		[	1			EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN
		1		<u> </u>		FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS
	1					GGSSGWGLGAGEEAEGDEDENMVYYFLEES
		}				TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
				!		LVIFKREKELARKLEFDGLYITEQPEDDDVKG
				]		QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG
		l				DIYGRERIAELLGMDLATLEITAHNERKPNPP
			İ			PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
						WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
	Ì					RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
		ĺ				AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL
		Í			(	FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV
					,	FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
						KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
						HNLANYMFFLMYLINKDETEHTGQESYVWK
						MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	Α	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
	ļ	ļ				LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV
<u> </u>	1050	<u> </u>	0000			N .
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD
	l	1	ł	}		YRHAP\PLLTNF\*FLVEMGFCYVGQAGRKLL
	i					ASSDQSALASQSAGITGISTAPGPPFFFLNFEA
503	1853	<del> </del>	3891	1772	1102	GSCSVAQAGVQ
202	1023	A	2071	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR
	[					QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP
	1					HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET
					}	WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC
	1			-		APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
					1	L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR
	1007	~	50,50		~	TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY
	l	l		1		QKEYYENQ*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFOKKVOFG
				-	.5.5	NENTKLELRKVPPELNNISKLNEHFSRFGTLV
	]					NLQVAYNGDPEGALIQFATYEEAKKAISSTEA
					!	VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL
	1	1	}	1	1	VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS
	i		1	ł	1	ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT
			1	i		
		ļ				LLVSTSAVDNNEAOKKKOFALKI.OODVRKR
		t:	i :			LLVSTSAVDNNEAQKKKQEALKLQQDVRKR
			i			LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE
			i :			LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMOAGEEVTEL
						LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NC: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA VITFKTRAEAEAAAVHGARFKGQDLKLAWN KPVTNISAVETEEVEPDEEEQREIIIA
506	1856	A	3911	1952	919	DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKORDKRNRHLGR
508	1858	A	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1004	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP MAALILVADNAGGSHASKDANQVHSTTRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHIHITWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHIHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS\*SQNPCSSPLFHHGL*AWLWCPELLLQGQARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863	A	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN VELGITISSPSSRKPGDPLVISDIKKGSVAHRT GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR YGGPLGVITISGTEEP\FDL*IISSLTKGGLAERT GAIHIGDRIL\AINSSSLKGKPLSEAIHLLQMAG ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSA\IDTS\YGTEGT\SFQASGY\NFNTYD WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED WDRSTASGFAGAA\DSAETEQEENFWSQALE DLETCGQSGILRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	A	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC
515	1865	A	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF\*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	A	3977	2	1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV
517	1867	A	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
-519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

	1 000 m	1 3 6 -4	1.050	l n - 1'-i - i		T A C-Countries
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	İ	in	i .	location	
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first amino acid		Q=Glutamine, R=Arginine, S=Serine,
ľ	İ	i			of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	Į	į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
			ļ	sequence		nucleotide insertion
		ļ '			<b>;</b>	LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
	<u> </u>	<u> </u>			100	AHV\FADLLLIITLPSYYIPFC
520	1870	Α	3999	882	698	QSFRLSLLSSWDYRHM*PRLANF*T\FFCRDR/
	<u> </u>					SLALLPRLVSNSWPQAILPPRPPKVLGLQT
521	1871	Α	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
						PPTSASHVAGATGTHHHAWLLSV
522	1872	A	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
		ŀ				IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
		ļ		ļ	}	EYGPVYSTWSALEGELAEPLEGVSACIGNCST
	<u> </u>	<u> </u>		L		AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
,		1		1	:	KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
ľ	1	ĺ	1	1	ĺ	L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
]	İ	1	İ			GNEGDITSFPAK
524	1874	A ·	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
	ľ	ł		}	1	TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
		1				RQGFTVLARMVLIS*PHDLPASASQSAGITGL
ļ		1	1	1		SHCSWPTSSILS
525	1875	A	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
						PPRLKQ/F/SHLSPPSIWDYRRVPPCLVNFSIFF
		1	1			VETGSCQPCLQLLGSSNPPASASQSAGIAGISH
1	ł	1	1	}		QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
		1			1	NKIINRPTHPVESSF
526	1876	A	4024	80	341	TPSSTSRGTEFQQSSKMAWQRREEKEHLNVR
	10.0	1	'•			RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ
1	ļ					LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	A	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
1	,	''	1025			DVAVYFTTKEWAIMG\PAERALYRDVMLEN
						YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
						TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	A	4028	1160	242	GTSELLCIORWNWGPAFPPRPGLALAPTLOLL
	1075	1	1,020	1		VEMGSAKSVPVTPARPPPHNKHLARVADPRS
	1	i	i	İ	l	PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ
		1				DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE
1	1	1		İ	,	VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
1		1				QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
1		1	1			ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
J	j	1			Į.	NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
1	ł		l		ļ	AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
1	1			1		WEQGQD\HDKENQHFPLVES
529	1879	A	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
1 329	10/3	1	5007	-	300	CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T
	i	1				ILTIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
[			1	ļ	1	HTEICT*MFIAVLFVVVKTWKQF
520	1000	<del>                                     </del>	4007	250	1,	
530	1880	A	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
1				1	1	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
	[	1	1		1	/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
	<u> </u>	<del>  </del>	<u> </u>			VKISILPKVNYRFYLISIKIIMAI
531	1881	Α	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY
		İ				HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
			ļ	1		IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
	<u>L</u>	<u> </u>	L		l	T*KR
532	1882	Α	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
1	1	1	1	1	1	YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
	1	1	1			KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL
		1				QIWMPVSLMNIVTLKCPT
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/
	Į	1				ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
						KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC
	<del>'</del>				<del></del>	·

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	)	09/496 914	correspondi ng to first	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	•			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						IFNAIPIKMPMMCMAKIEKNSS
534	1884	Α	4088	3	1931	IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC
]		]	}			GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL
						QTRLVDAAKALNLVHCHCLDIFINQAFDMQR
						DLQITPKRLEYTRKKENELYESLMNIANRKQE EMKDMIVETLNTMKEELLDDATNMEFKDVI
			l .			VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
<u> </u>			,			NKLISSVDYLRESFVGTLERCLQSLEKSQDVS
						VHITSNYLKQILNAAYHVEVTFHSGSSVTRM
			!			LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI
						ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ
		ľ	1			LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS
						LESRSLQDVLLHRKPKLGQELGRGQYGVVYL
		•				CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA
			}			VLLIMERLHRDLYTGLKAGLTLETRLQIALDV
		·	Í			VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI
ļ.						TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK
						YDNSVDVYAFGILFWYICSGSVKLPEAFERCA
						SKDHLWNNVRRGARPERLPVFDEECWQLME
						ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\
535	1885	A	4090	2	417	NSEQPNRGLDDST
333	1003	^	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
	<u> </u>					HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR
ĺ			[			ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK
						HNRKRIWLRA
536	1886	Α	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI
						PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK
627	1007	<u> </u>	4104			EQNLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML
			]			IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE
538	1888	A	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSOPL
		' '	1			VPRENITAWMNAIGLIITALPVS
539	1889	A	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM
	ŀ					I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
						PLFPKIYREGTLPHSFYEASITL
540	1890	Α	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE
						AALGDFLGLHRRTQQPAVDRLLSDASAQWR
		1				VRGHGGVRESGRAPQQPGRRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN
1	1	Ì				DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH
[			]			IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR
						CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT
						MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY
	[				•	GFPELEDSTSEYLCTILNIQNVCMTFDVASLY
		]				SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK .
1	1	Į	1			TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK
[	1	1	ſ	·		ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT
			<b>!</b>			MKYGAQVVKGELKSALLDGDTQNYDLDHG
			1			FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR
		}				DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS
l	ľ	Ì	ľ			WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF
	ļ		ł			ECMFTNKTFTLEKGLIVPMENVATIADCASVI
		'	1		-	EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
541	1901	<u> </u>	4146	202	770	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY
741	1891	<u> </u>	4146	282	778	GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

					<del></del>	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 201100	ļ	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i	i			sequence	
		1	1	peptide	1	/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
	ļ	ļ	l		}	HAESENFAFWQDMKWKNKFWGKSLEIVPVG
i		l				TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
1		l				IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
1	1	ļ				VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
}		l	1			PTPGGOCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA
342	1092	^	414/	44	433	
i						QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR
l .		i	ļ			INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT
		1	1			ANVSSENNTPRTSKTTFQLELSVKDAVYTVV
1		Į	Į.		}	SSH
543	1893	A	4153	678	11	TISYPOCLTOMYFLISFANVDTFLLPIMALDH
		1	I			YVAICSALQ*CSIITP/ELCQGLPVLA*AGSSLIS
		1				PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA
		1	1	•		
ł	1	ł	1	ł	1	CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA
		1	1			AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV
1			j			LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFTY
	1					SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	Α	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP
1	l	i	l	l		SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL
1	1	I	1			LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS
ł	i		İ			LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG
ı			į			
			l			LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE
L		ļ	L			LKVGREGHVLPWQAHVVEF
545	1895	Α	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
	i					DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
						QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
1						ELELSGEPAPGWQVLAGYTYTQARYLRDASE
	i	1	l	ł	Î	ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL
340	1050	^	71/7	1222	1150	OPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
1		1	1			
		i	1	1	1	FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA
	L		L		<u> </u>	GITGVSHHA*PRASGRRCW
547	1897	Α	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP
1		1	Į.			GSHHASEGLAPGVPPAGGVSAQELTAPPQEG
	1		l .		1	WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA
	ļ			!		RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS
1	1	1	[		ĺ	GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ
1	1		1			PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL
ŀ		1		1		AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP
		I				HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS
1		1				TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP
		1		ļ	]	ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/
		1				RCPPACSPAAASSFSFESQPCPSAPSKASPAPA
1		1		1		AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL
1	1	}	}	1	1	AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA
		1				LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/
	1	I				PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
			1			1
1		1			İ	AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ
į	ļ	Į.	ļ			ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
1		1	1	1		PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
1		ļ			ł	RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC
1	i	1	1	1		LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS
1		i		1		WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL
1	1	ŀ		1	)	PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ
		I		1	1	PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL
		ľ		•	•	
1		1				TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
I		1		]	]	MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP
1	Į.	!	Į.	1	1	P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT
<u> </u>		<u> </u>	<u></u>		i	PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG LHLPGGRTK
548	1898	A	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVFHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQELKL LQEQDQELCEILCMPHYDIDSASVPSLEELNQ FRQHVTTLRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATL QKLLRQULEMQKSQNEAVCEGLRTQIRELW DRLQIPEEEREAVATIMSGSKAKVRK\ALQULE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEEKQRAKLQKMLP KLEEELKARIELWEQEHSKAFMVNGQKFME YVAEQWEHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLL\ICITVCLSYLPE AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSITFP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO:	nucleotide	location	
eotide	seq-	1	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	
uence	1 401100		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
Helice		}	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ	ĺ			residue of		
					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
552	1902	A	4197	sequence 2	14302	
332	1902	A.	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR
						EPAPARRRGTMADGGEGEDEIQFLRTDDEVV
1			ł			LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
1			l			NSKNVPPDLSICTFVLEQSLSVRALQEMLANT
i						VEKSEGQVDVEKWKFMMKTAQGGGHRTLL
1			}			YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
						VGLQEDTTGEACWWTIHPASKQRSEGEKVR
{						VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
						QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
						GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
1						VHARSLWRLETLRVAWSGSHIRWGQPFRLR
1						HVTTGKYLSLMEDKNLLLMDKEKADVKSTA
1						FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
1	i l					VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
						KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
						TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
1						SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
i l						VAGREAGESWKSILNSLYELLAALIRGNRKN
1 :						CAQFSGSLDWLISRLERLEASSGILEVLHCVL
	}					VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD
						VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL
1			ľ	' i		LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
1						YELMVDHTEPFVTAEATHLRVGWASTEGYSP
						YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG
1 1	1					CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
1			1			RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
						RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
1	ł			1		KVEHSREYKQERTYTRDLLGPTVSLTQAAFT
						PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
ľ				ľ		KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
						ERNYNLOMSLETLKTLLALGCHVGISDEHAE
1						DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
	1	1				PSQEAMVDKLAENAHNVWARDRIRQGWTY
	1					GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS
						LREAVRTLLGYGYNLEAPDODHAARAEVCS
		ł	l	ł	ł	GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
		ļ	1		į	GDMRVGWSRPGCQPDQELGSDERAFAFDGF
		İ			ļ	KAQRWHQGNEHYGRSWQAGDVVGCMVDM
	1	1			l	NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
					ļ	GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC
	Í	Í	ĺ	j	ſ	GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV
						PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN
		ļ	1	•		SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
1	1	i			ľ	LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
}				l	İ	DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
	ļ	]	1	j		RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
	1	- 1	ľ	ì		DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
	1	1		1	İ	DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE
Į į	İ	1	ļ	l		SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
1	}	1	l		ł	VVDAASGLLTFIANGKELSTYYQVEPSTKLFP
	i	1	ļ	1		AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
	Í	ĺ	[	ĺ		EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
	İ	ł	ŀ			KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
		1	ļ			RSVDILELTEQEELLKFHYHTLRLYSAVCALG
	. 1	i	į	ľ		NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
	-	1			ļ	AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
			į	1		EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF
		ł	1	- 1	İ	SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE
L						AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
				soquenec		MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
						LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL   QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
						VAFSDDFVAKLQDNQRFRYNEVMQALNMSA ALTARKTKEFRSPPOEOINMLLNFKDDKSECP
						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
					:	NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES DSKKSSTLQQLISETMVRWAQESVIEDPELVR
						AMFVLLHRQYDGIGGLVRALPKTYTINGVSV EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
		!				LGDIMNNKVFYQHPNLMRALGMHETVMEV MVNVLGGGESKEITFPKMVANCCRFLCYFCR
						ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
						STPLDVAAASVMDNNELALALREPDLEKVVR YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
						YLDFLRFAVFCNGESVEENANVVVRLLIRRPE CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
ļ						GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
						LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
ł						AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL EVGFLPDLRAAASLDTAALSATDMALALNRY
						LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
1						MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH YERCWKYYCLPGGWGNFGAASEEELHLSRK
						LFWGIFDALSQKKYEQELFKLALPCLSAVAG
}						ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
}						LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM
						ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
			•			NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
						QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
						PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
						RISLFGNDATSIVNCLHILGQTLDARTVMKTG
						LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
						GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD
						KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
						LCSYMSRWWEHGPENNPERAEMCCTALNSE
						HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
						SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
						MVAEVFIYWSKSHNFKREEQNFVVQNEINN
						MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
						LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL
						DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR
				•	- 	HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
						KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						EEEVKSFEEKEMEKQKLLYQQARLHDRGAA EMVLQTISASKGETGPMVAATLKLGIAILNGG NSTVQQKMLDYLKEKKDVGFFQSLAGLMQS CSVLDLNAFERQNKAEGLGMVTEEGSGEKV LQDDEFTCDLFFFLQLLCEGHNSDFQNYLRT QTGNNTTVNIIISTVDYLLRVQESISDFYWYY SGKDVIDEQGQRNFSKAIQVAKQVFNTLTEYI QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVESSNNVEMIL KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK RDFHKAMESHKHYTQSETEFLLSCAETDENE TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG SAKRIERVYFEISESSRTQWEKPQVKESKRQFI FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEEQGPRMAFF SILTVRSALFALRYNILTLMRMLSLKSLKKQM KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEEGERKPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW
553	1903	A	4199	31	767	KLGVVFTDNSFLYLAWYMT  LPELNGRGAGLRRAEPSERGGAERTQQVAA LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQIFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS
554	1904	A	4200	1	961	GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT\ RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH
555	1905	A	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNP NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLS DSNRDHTANRQQRST\SPVARRTRSQTSVNFN GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

NO: of nucle-citide cotide unce collected country of the control o	C050 110	CEO ID	1 1/	000	1 15 12 1 1	N - 31 - 4 - 3 4	1 4
nuclecule seq uence USSN location corresponding to the properties of the properties	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence		ı	nod				
uence    9/146   09/496   0   0   0   0   0   0   0   0   0							
uence   914   ng to first amino acid residue of peptide peptide peptide sequence							
amino acid recidue of peptide sequence propried	•	nence	1	1			
residue of peptide sequence Profession x-Vossible nucleotide deletion x-possible nucleotide deletion x-possible nucleotide insertion  GGAGGPRANASRYINFSSHTNOSGGSELRQRE GQRFGAAHVWENGARSNYTVENTHORLEF! RLRSTSNSRSRSPIGOSGGTVAHNSQRESRPV QQTTRRSVRRGRTRVFLEQDRERERGTAY TFFSNSRLVSRITVEGEGESSRSSTA XRHPHTI LDLQVRRIRRGENGSGTVAHNSQRESRPV QQTTRRSVRRGRTRVFLEQDRERERGTAY TFFSNSRLVSRITVEGEGESSRSSTA XRHPHTI LDLQVRRIRRGENGEDISANKTESRVGLAE NVTUENSGGFRGTISHERGGRITVSTITVP LRRISENGLVPSSVALRSLRQIMTGFGELSSL MEADSESLQRNGQHLPPMISELSHLGTDN NRSQHREGSSQDRQAQQDSTEMHGENETTOP HERNSDRGGRGTQLRPMINLVETTLE RLAH FFILLNESDDDDRRGLTKEQDDNLSTRHYEHN SISSELGKICSVCISDYVTORLRQLCMHEF HHCIDRWLSENCTCPICRQPVLGSNANNG LORGREGASCHEVYCVTORLRQLCMHEF HHCIDRWLSENCTCPICRQPVLGSNANNG SISSEPTLSWCGGFFAAAPVNLVFTTLTRLAH FFILLNESDDDDRRGLTKEQDDNLSTRHYEHVON SISSEPTLSWCGGFFAAAPVNLVFTTLTRLAH FFILLNESDDDDRRGLTKEQDDNLSTRHYEHVON SISSEPTLSWCGGFFAAAPVNLVFTTLTRLAH FFILLNESDDDRRGGRQLRAHSSTANNST SISSEPTLSWCGGFFAAAFVAGKXEER QRAGKEGTAPSENGETKABEHHISSSTANNST SISSEPTLSWCGGFFAAAFVAGKXEER QRAGKEGTAPSENGETKABEHHISSSTANNST SISSEPTLSWCGGFFAAAFVFLVET GFILLNOGAGHKLPTSGDPASASQARTIOMS HRTWFLASFLIDSCKNFTVYKUMYTL TYPHAFEERHFFTSSKATTVSYDYTRIRGFRKLID GFILLNOGAGHKLPTSGARTYSFTVAGAGKABE LDCCFNTTORGRETOQGVREPKFPSKKDSLAC TYSNKNDVDLRSSVDKWREKKREGGCCNE SNSSSNQLDKSSGKLPVKRSTNNLKKSJTVK VPVKTYDTRTA ASSTREQMSHDLVAV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGRETOQGVREPKFPSKKDSLAC TYSNKNDVDLRSSVDKWREKKREGGCCNE SNSSSNQLDKSSFSKNYTTQRKDITHHEFTSKMADHLAV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLVQ KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLV KSLTICASTORGTCTPROGTTHKFREELMRK KVV VPVKTYDTRTAHSTRGRCCPYNKSGLV KSLTICASTORGTCTPROGTTHKFREELMRK KV	uence	}	l	914			T-Thronia W. W. II.
peptide sequence   Apostòlic nucleotide deteiton, '-possible nucleotide insertion   nucleotide insertion   GGAAGIPRANASR'INFSSHTNOSGGSELRORE   GGRAGAHVWENGARSN'IVVENTIVER, PER   RIRSTSNSSSSSPIOROGGIVYHNSQRESERPV   QOTTRESVERRORTRYTLEONEGERERGTAY   THE SYRERORTRY THE SYRERORTRY THE SYRERORTRY THE SYRERORTRY THE SYRERORTRY THE SYRERORTRY WITH   LICHOVERRIFGENERD STAN ARRHPHTI   LICHOVERRIFGENERD STAN ARRHPHTI   LICHOVERRIFGENERD STAN ARRHPHTI   LICHOVERRIFGENERD STAN ARRHPHTI   LICHOVERRIFGENERD STAN ARRHPHTI   LICHOVERRIFGENER STAN ARRHPHTI   LICHOVERRIFGENER STAN ARRHPHTI   LICHOVERRIFGENER STAN ARRHPHTI   LICHOVERRIFGENER STAN ARRHPHTI   LICHOVERRIFGENER STAN ARRHPHTI   LICHOVER STAN					1		
						sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
GGAAGIPRANASRTNYSSHTNOSGGSELRÖBE GGRFAAHYWENGARSNYTVRINTORGESIRY QGTREASYRNGRISTRYLEGNEGRESRGTAY TFSNSRLSSPIOROGGIVYHNSQRSSRYV QGTTRESVERGGRTGVELGDEGSRSSTA ARRHPTIT LDLQVRIRRGPENDRDSIAMRIESSE VGLAE NTVTIESNSGGFRITISRLERSGIRTYVSTITVP LERISERLEVPSSVALARSILRQMINGFGELSSL MEADESSELORIGQHLPDMISSL SNLGTDN NSQNIRGGSSDROAQODSTEMHEGHENTTOP HTRINSDRGGROLRAPMINUTGTTLPLLAH FELLINESDDDDRIRGLTKEQDUM. STRIYLEIN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HHICIDRWLSENCTCPICRPYLGSNAMNNG THE HHICIDRWLSENCTCPICRPYLGSNAMNNG SUPERINGROSSORY ARREST STRIYLEIN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HHICIDRWLSENCTCPICRPYLGSNAMNNG KSPENTEGKDOSK VTKQETTRENARLSAKPA RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK RKPENTEGKDOSK VTKGCK RKPENTEGKDOSK RKPENTEGKDOSK VTKGCK RKPENTE							
GQRFGAAHVWENGARSNYTVENTHORLEPI   RIATISNSKSRSPIORGSTVHNSQRSSPY   QQTTRRSVRRGRTRVFLEQDREKERGTAY   TFFSNSKLVSRITVEGEGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTA ARRHPITI   LOLQVERIRGERGESSSSTAR ARRHPITI   LOLQVERIRGERGESSSSTAR LESGRITYVSTITVP   LRRISENELYPESVALRSILROIMTGFGELSSL   MEADSESLQRNQQHLDPMSHSLSNLGTDN   NRSQHREGSSQDRQAQGDSTEMHEGRETICP   HTRNSDSRGGRQLNPPNNLVETGTLPILLALH   FELNESDDDDRIRGLTEQEDNLSTRYPHN   SIDSELGKICSVCISDYVTGNKLRQLCCMHEF   HHCIDRWLSENTCTICRCPPLICSNLANNG   LORQRQHPAAAPAYPVRCTFCFTDIVIMPKR   SPENTERSKASCHAPSNCGTERKAEHLINSSTANYNST   SRGTPPSTLSVKQGIETVRVKGTEN ARRHPITE     SPENTERSKASCHAPSNCGGETKAEHLINSSTANYNST   SRGTPPSTLSVKQGIETVRVKGTEN ARRHPITE     SPENTERSKASCHAPSNCGGETKAEHLINSSTANNST   SRGTPPSTLSVKQGIETVRVKGTEN ARRHPITE     SPENTERSKASCHAPSNCGGETKAEHLINSSTANNST     SRGTPPSTLSVKQGIETVRVKGTEN     SRGTPPSTLSVKQGIETVRVKGTEN     SRGTPPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKQGIETVRVKGTEN     SRGTPSTLSVKGGIETVRVKGTEN     SRGTPSTLSVKGGIETVRVKGTEN     SRGTPSTLSVKGGIETVRVKGTEN     SRGTPSTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTEN     SRGTPSTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEHLINGTLSVKGGIETGLAEH					sequence		
R.R.STSINSRSSPJQQGGTTVYHNSQEESPR			ļ				
See		ĺ	1		1		GQRFGAAHVWENGARSNVTVRNTNQRLEPI
					Ì		RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV
LDLQVPRIRPGENRIDEJSIARTRISRVCIES		l					QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
					i		
LRISENELVEPSSVALRSILRQIMTGEGELSSL MEADSESELQENGOQH DPMGNETTOP HTRNSDSRGRQDRQAQGDSTEMHGENETTOP HTRNSDSRGRQDRQAQGDSTEMHGENETTOP HTRNSDSRGRQDRQAQGDSTEMHGENETTOP HTRNSDSRGRQLRNPNNLVETGTLPLLALH FFLLNESDDDDRRGLTKEQDINLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HHLCDRMLSENCTCPICRQFULGSNIANNG LQRQRQIPAAAPAVPVRCFTFCFTDVIMFKR KSPENTEGKDGSKVTKQEPTRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKEKEK QEAGKEGTAPSENGETKAEBLISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN  557 1907 A 4213 774 507 ARRSSCLTQTSWGGHENGPRANDVFLVET GFLHIQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLDSCKNFIVYKIMFTL GDQDFSDGRTQKYCKEEDRXYSFQKGPINE LDCFNTGRGRETQDQVKEEDRXYSFQKGPINE LDCFNTGRGRETQDQVKEEDRXYSFQKGPINE LDCFNTGRGRETQDQVKEEDRXYSFQKGPINE LDCFNTGRGRETQDQVKEPKPSKKDSIAC TYSNKJDVDLRSSNDKWKEKKKKEGDCRE SNSSNQLDSKQLPDVKPSPKPSKKDSIAC TYSNKJDVDLRSSNDKWKEKKKKEGDCRE SNSSNQLDSKQLPDVKPSPKFSKKDSIAC TYSNKJDVDLRSSNDKWKEKKKEGDCRE SNSSQLAGSKQLPDVKPSPKFSKKDSIAC TYSNKJDVDLRSSNDKWEEKKKEGDCRE SNSSQLAGSKQLPDVKPSPKFSKKDSIAC TYSNKJDVLRSSNTTQRODITHFRFEREMMR KKVP  SSSQLAGSKQLPDVKPSPKSKDSAIC RSSSNQLDSKQLPDVKPSPKPSKKDSIAC TYSNKJDVLRSSNTTQRODITHFRFEREMMR KKVP GRIRNTRVPFKSNFRGGRGPNYKSGLVQ KSLYJQAKYQRLRTTGGRGFTHKFRERLMK KKVP  SSSQLAGSKAVTVPKPROPTITGLFGFLEFT FFHHVGQAGLELLTSGDPPALAFQSAGTTGYS HAWLQVLNS  1560 1910 A 4246 2 1569 TLSLLERVLMDDIVTPYPQEEVKTVIRKCLEQ AALVNYSSLSSPAKECKKNPTLPPFCLA SQVMDLTIQNQKDAENVGRHTPAKKLEDTR LAELVEVLQQNEEHHAEAFAWWSDLMVEH AETFISLFAVMDDAALEVQPDTWDSFFLFG HLEQRIKLMASMMESCVKRENGLERSESWEPVNN GSGTSSEDLFWKLDALQTTRDLHWPEFGK HLEQRIKLMASMMESCVKRENGLESRESWEPVNN GSGTSSEDLFWKLDALQTTRDLHWFEFGK HLEQRIKLMASMMESCVKRENGLESRESWEPVNN GSGTSSEDLFWKLDALQTTRDLHWFEFGK HLEQRIKLMASMMESCVKRENGLESRESWEPVNN GSGTSSEDLFWKLDALQTTRDLHWFEFGK HLEQRIKLMASMMESCVKRENGLESRESWEPVNN GSGTSSEDLFWKLDALQTTRDLHWFGETJEFG KEMITLVAKFYTILEGVLAKLSRYDGETLFS SFLSTVKAASKYVDVPKPGMDVADAAVYTF VRHSQDVLRDKKVPTLEGVLAKLSRYDGETLFS SFLSTVKAASKYVDVPKPGMDVADAAVYTF VRHSQDVLRDKKVPTLEGVLAKLSRYDGETLFS SFLSTVKAASKYVDVPKPGMDVADAAVYTF VRHSQDVLRDKKVPTLEGDLKHVTGGDYKDI		1 -	ľ	i	ĺ		LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE
MEADSESELQRIGGHLPDMHSELSINLGTDN   NRSONREGSODROAQGOSDROMEGNET (SUME CHIEF LIR LAH FELLNES)   FELLNES DIDDERIG LTKE QUDNLS TRHYJEHN   SIDSELGKICS VCISDY VTONKLR QLPCMHEF   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   SIDSELGKICS VCISDY VTONKLR QLPCMHEF   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   CARCAMER   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   CARCAMER   CARCA							NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP
MEADSESELQRIGGHLPDMHSELSINLGTDN   NRSONREGSODROAQGOSDROMEGNET (SUME CHIEF LIR LAH FELLNES)   FELLNES DIDDERIG LTKE QUDNLS TRHYJEHN   SIDSELGKICS VCISDY VTONKLR QLPCMHEF   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   SIDSELGKICS VCISDY VTONKLR QLPCMHEF   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   CARCAMER   HIHCIDR MLSENCT CIPIC (RV VLOSMANNG   CARCAMER   CARCA							LRRISENELVEPSSVALRSILROIMTGFGELSSL
		ļ					
HTRNSDSRGRQLRPPNNLVETGT_LEILALAH		Ì	ļ	ļ			
SIDSELGKICSVCISDYVTGNIKIRQIPCMHEE		ł	l	l	ł		
		İ					SIDSELGKICSVCISDYVTGNKLROLPCMHEE
1906		i		1			
S57 1907 A 4213 774 507 ARPFSCLTQSGKVTKQEPTRRSARLSAKPA PPKPEPKPKTSAKKEPGALISAGGAGKKEED SGRIGHT ABENDRISTYNVST SRGTPFSTLSVKGQIETVARVGTEN ARPSSCLTQTSWGRIFKGPFRANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNTTVYKIMYTL.  558 1908 A 4225 3 1253 TYRHAEREHPETSSATKVSYDYRHKRPFKLD GDQDFSDGRTQKYCKEEDRKYSPGKGPLNRE LDCPNTGGRGETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT HIQVKAANYPPSSGITHERSSKMADHHADV NEPHNSDPEHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRRKGRLQNEDEHTPH ASAAERDDQNSSFSKNYTTQRKDILTHKPFEV EGNHRNTRVRPFKSNFGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGGRGTQHTGKDHTHFLFFS KAADHKADV NEPHNSDPEHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRRKGRLQNEDEHTPH ASAAERDDQNSSFSKNYTTQRKDILTHKFPEV EGNHRNTRVRPFKSNFGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGGRGTIDHHTGLIFGFLIET EFFHYQQAGLELLTSGDPPALAFQSAGTIGVS HAWLQVLNS  559 1909 A 4235 1 323 KFSIPFTLWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFFHYQQAGLELLTSGDPPALAFQSAGTIGVS HAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVPCL AALVNYSRLSEYAKIEGKKREMYELPVPCL LLINDFLRTGLLLICONGKVFHKHLQDLFAPLVV RYMWDLOSSPIAQSHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHVPEEFFCK HLEQRLKLMASDMIESCVKTRILAFEVKLQK SIQQVLDSTLYNKLGLALQTIFRDLHVPEEFFCK HLEQRLKLMASDMIESCVKTRILAFEVKLQK SISQQVRPVQFFNAMGAGSIQP KLVCSMEMGQEFAKMWQYHSKIDELIEETV KEMITLLVARFVTLEGVTAKLSRYDEGTLFS SISJSFTVKAASKYVDVPKPGMDVADAYVTT VRHSQDVLRDKVNEEMYTERLTQWYNSSM NICTWLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQLHIVQLKTLIRMVKTY RDFRLQVLLTDRMDLQHIVALSTLARD SVSEGGGLQGISMKDDSEEDEEDD	556	1906	Δ	4212	1 2	462	
PPKPEPKPRKTSAKKEPĞAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKOGIETVRVKGTEN	330	1200	1 ^	4212	,	402	
1907   A   4213   774   507   ARRESCLTLQTSWGRIENGRASSASARITGM		İ		1			
SRGTPPSTLSVKGQIETVRVKGTEN							
1907   A   4213   774   507   ARRESCLTLQTSWGHRENGPPREANFVELVET GEHIGQAGHKLPTSGDPPASASQSARITGMS   HRTWFLASFLIDSCKNFIVYKIMYTL.							CEAGREGIAPSENGETRAEEIHISKSIVNVSI
GFLHIGQAGHKLPTSGDPPASASQSARITGMS   HRTWFLASFLIDSCKNFIVYKIMYTI.	5.50	1007		4010	774	505	
HRTWFLASFLIDSCKNFIVYKIMYTL    S58	337	1907	A	4213	//4	307	
1908   A   4225   3   1253   TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQGPLNRE LDCPNTGRGRETQDGQVKEPFRSKKDSIAC TYSNKNDVDLRSSNDK WEKKKKEGDCRKE SNSSSNDLKSQKLPDVKPSPNLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFIBPVFEHLDSTQNTENRPTGEFAQBIIT IIHQVKANYFPSPGITLHERFSKMADHKADV NEPPLNSDPEHRRIDMSLAELQSKQAVIVESE QTLIKIDDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKFFEV EGNHRNTR VPFFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFFERLMRK KKVP							
GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDQVKEFFKPSKDS1AC TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFIFPYEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGTILHERFSKMADIHKADV NEIPLNSDPEHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKFFEV EGNHRNTRVRPFKSNFRGGRCQFNYKSGLVQ KSLYIQAKYQRLRFTGPRGFTTHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFILRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLIS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSYAKIEGKKREMYELPVFCLA SQVMDLTJONQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLVNDFLRTGLLICGNGKVHKHLQDLFAPLVV RYMWDLDGSSPIAQSHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTTIRDLHWPEEEFGK HLEGRLKLMADMIESCVKKTRILAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGIMAKGSIQP KLCSMEMGQEFAKWHQYHSKLIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHTVQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNBLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPJKRADILKHVIGDYKDI		1000	L		<u> </u>		
LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFSKMADHIKADV NEPLNSDPEHHRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAAERDDQNSSFSKNYTTQRKDIITHKFPEV EGNHRNTRVRPFFSSFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFREIMRK KKVP  559 1909 A 4235 1 323 KFSIPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGI_ELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEFGK HLEQRLKLMASDMIESCVKRTRUAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLICSMEMGGEFAKMWHQYHSKIDELIETTV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHTVQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNBLTVEEATA SVSEGGGLQGISMKDDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKPIKRADILKHVIGDYKDI	558	1908	A	4225	3	1253	
TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFIPVFEHLDSTQNTENKPTGEFAQBIIT IIHQVKANYFPSPGITLHERFSKMADIHKADV NEIPLNSDPEHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFILVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEHHAAEAFAWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLINDFLRTGLLICGNGKFHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKTTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLNCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMFJELFQVLYNSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLINSKTYTETIRNRLTVEEATA SVSEGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		l					GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE
SNSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFIPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSFGITLHERFSVKMADIHKADV NEPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQFNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELFVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLNDFLRTGLLICGNGKWFHKHLQDLFAPLVV RYMWDLDGSSPIAQSHIRGLISRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPGFNMAPCFNVMGLMAKGSIQP KLNCSMEMGEFAKMWYQHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMHYENLFDQWYNSSM NVICTWLTDRMDLQHHYQLKTLIRMVKKTY RDFRLQGVLDSTLINSKTYTETIRNRLTVEEATA SVSEGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPKRADILKHVIGDYKDI			ļ				LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC
VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEITI IIIHQVKANYFPSFGITLHERFSKMADHHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKREMYELPYFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTR LAELVIEV-LQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLINDFLRTGLLICGNGKFHKH.QDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRIKLMASDMIESCVKTRRIAFEVKLQK TSSIQIFRVPGFNMAPCFNVMGLMAKGSIQP KLICSMEMQGEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLINSKTYETIRNRLTVEEATA SVSEGGLQGISMKDSDEEDEEDDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI			i		· '		TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE
RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSGTTLHERFSVKMADIHKADV NEPILNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLIMDFLRTGLLICGNGKYHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRILAFEVKLQK FSSIQQIFRVPQFNMAPCTNVMGLMAKGSIQP KL\CSMEMGQEFAKMWQYHSKIDELIEETV KEMITLLVAKFVTILLEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLINSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPKRADILKHVIGDYKDI			1				SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK
IIHQVKANYFPSPGITLHERFSKMADIHKADV NEIPLNSDFEIHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAERDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSSPLFQ LL\NDFLRTGLLICGNGKVHKHLQDLFAPLVV RYMWDLDGSSPLAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMESCVKRTRILAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLINSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							VDVKKTVDTFRVASSYSTERQMSHDLVAVG
NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNIRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  1909 A 4235 1 323 KFSIPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLNDFLRTGLLICGNGKVHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRUAFEVKLQK TSSIQQIFRVPQFFNMAPCFNVMGLMAKGSIQP KLVCSMEMGQEFAKMWHQYHSKLDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRITVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPKRADILKHVIGDYKDI							RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT
OTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRUAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLVCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							IIHQVKANYFPSPGITLHERFS\KMADIHKADV
OTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRUAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLVCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI					ŀ		NEIPLNSDPEIHRRIDMSLAELOSKOAVIYESE
ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNITRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLITIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLNDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYMWDLDGSSPLAQSHRGGLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRUAFEVKLOK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLVCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		l .					
EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP  559 1909 A 4235 1 323 KFSIPFFLRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV RVYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI					İ		
SSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP							
S59   1909   A   4235   1   323   KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS					}		
1909   A   4235   1   323   KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS     560   1910   A   4246   2   1569   TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\nDFLRTGLLICGNGK\FHKHLQDLFAPLVV R\mathcal{YMWDLDGSSPIAQSIHRGLLSRESWEPVNN} GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSS\iQQ\iff VPQ\text{FINAPEFTVMGLMAKGS\iQQ KL\CSMEMGQEFAKMWHQ\mathcal{YHSK}IDEL\ieeTv KEMITLLVAKF\mathcal{YTILEGVLAKLSR\mathcal{YDEGTLFS} S\text{FLSFTVKAASK\mathcal{YVDVFR}GMD\ndayvTF} VRHS VRHS VVTF VRHS VRHS VVTF VRHS VNTEMYUTDRMDLQLH\mathcal{YUTDR					ì		
LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560  1910  A 4246  2 1569  TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGKYFHKHLQDLFAPLVV RYYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEFGK HLEQRLKLMASDMIESCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLVCSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561  1911  A 4257  1300  654  SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	559	1909	A	4235	1	323	
EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS  560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGK/FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR/LAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL/CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI					1		
HHAWLQVLNS				1	1		
1910   A   4246   2   1569   TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ   AALVNYSRLSEYAKIEGKKREMYELPVFCLA   SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR   LAELVIEVLQQNEEHHAEAFAWWSDLMVEH   AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ   LL\nDflrtgllicGngk\fhkhlqdlfaplvv   R\gammaymWdldgsspiaqsihrgllsreswepvnn   Gsgtsedlfwkldalqtfirdlhwpeeefgk   HLEQRLKLMASDMIESCVKRTR\iAfevklqk   TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP   KL\csmemgqefakmwhqyhskidelieetv   Kemitllvakfvtilegvlaklsrydegtlfs   SFLSftvkaaskyvdvpkpgmdvadayvtf   Vrhsqdvlrdkvneemyierlfdqwynssm   nvictwltdrmdlqlhiyqlktlirmvkkty   RDfrlqgvldstlnsktyetirnrltveeata   Svsegglqgismkdsdeedeedd   Selvqfllikdqkkipikradilkhvigdykdi		[ .	[	1			
AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R\YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	560	1910	_	4246	-	1560	
SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R\YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	500	****	^	1210	ا "	1307	
LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMÆSCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI			1				
AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		l	!	1	l	!	
LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMŒSCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI			l	1			LAEL VIE VLUQINEEMHAEAFAWWSDLMVEH
R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1	l .	1	Į.	<b>i</b>	AEITLSLIAVDMDAALEVQPPDTWDSFPLFQ
GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		l	l	1			
HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1	1		ĺ	ļ	WYMWDLDGSSPIAQSIHRGLLSRESWEPVNN
TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1	l	1			GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK
KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1	ĺ	ĺ	( ·		HLEQKLKLMASDMIESCVKRTR\IAFEVKLQK
KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		]	I				
SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI				1	1	i	KL\CSMEMGQEFAKMWHQYHSKIDELIEETV
SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1	İ		1		
VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		}	].	}	Į.	}	SFLSFTVKAASKYVDVPKPGMDVADAYVTF
NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		1					
RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD  561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI			i				
561         1911         A         4257         1300         654         SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI		l	1	l			RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA
561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI			ľ		1	-	
	561	1911	Α	4257	1300	654	
			l	1	l		FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI FMKGNTIKETEAWDFLLALIGVYPTKKHLIFG DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ WGPRTNI FTSKMKVI KEVAKVHNODPKDW
562	1912	A	4260	1	1498	WGPRTNLETSKMKVLKFVAKVHNQDPKDW PAQYCEALADEENRARPQPSGPAPSS MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF WLHARLQKCFLSRGCGSYCAGAKASPLPGK MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV ASGETADVVQTAAEQSFAELGLGSYTPVGLI QNILEFMHVDLGLPWWGAIAACTVFARCLIF PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA GDHIEYYKASSEMALYQKKHGIKLYKPLILPV TQAPIFISFFIALREMANLPVPSLQTGGLWWF QDLTVSDPIYILPLAVTATMWAVLELGAETG VQSSDLQWMRNVIRMMPLITLPITMHFPTAV FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR VVHDLDKLPPREGFLESFKKGWKNAEMTRQ LREREQRMRNQLELAARGPLRQTFTHNPLLQ PGKDNPPNIPSSISSSSSSKPKSKYPWHDTLG
563	1913	A	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF LIQENNNTNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF
564	1914	A	4270	3	368	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	A	4288	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GOTGASRTPRT
566	1916	A	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR
567	1917	A	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG
568	1918	Α	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY
569	1919	Α	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT
570	1000		4000		0.00	VTESKLEAEGKTKEKAREKERKKKS
570	1920	A	4308	3		RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMILLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYV\ETL\KHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLKÆRM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	A	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

COTO ID	I CEOTO	1 1 ( -4	LCEO	Dun di eta d	Deadisted and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	peptide	nou	in NO.	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciico	İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uchee		1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		·		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		J	}	peptide	Joquen	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
						ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
					}	FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
					ļ	SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
					1	SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
ł		l	ì	1	1	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
						CWPGWSSTPDLK
580	1930	Α	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
		1				\VFKKGI\IHILHELFQNKEEGAFPNS/FYEASFT
		1				LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
		J -				QLKSSDL
581	1931	A	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
		1				RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
ł						RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
1	[		[			RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
						LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
					1	VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
				ĺ		DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
						SPE
582	1932	Α	4424	194	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
l		1				LEQELLEHGRDAASVQAATSVQAMQGKTTL
	1					PS\QGPLQRPSRLVFT\DVANAIHV
583	1933	Α	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
						PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
	1024	<u> </u>	4400		(00	SAPPALLQDTSV
584	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
l				ļ		APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
ŀ						APATQHSQAGPATGQAYGPHTYTEPAKPKK GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
į				l		ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
İ		1	1			GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
		ļ			ļ	QGPHGKAAQGGAAGAAGRLGLYH
585	1935	A	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
1 303	1,555	ļ	11103	"	1	SIFDDFAHYEKRQ
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
			,			FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
					1	INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
			1	İ		TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
,			l	1		FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
				1		LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
	1	1	1	}	ŀ	PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
į.						ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
			1			LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
			1			PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
						LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
]	1		1		1	SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
		1		]		FRAPPAINARLPFNFFFPFV
587	1937	Α	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
			1	l		CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
	1	1	1	1		NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
			1			CPANFCIIII/DFLVETGFHHVGQASHELLTSGD
1			ŀ			PPTSASQSAGITGMSYHTWFGES
589	1939	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
	1		1			PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
						REAAQAMYIDSYNSRGFMINGNRVLGPCALL
	1					PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI
		-	•		<del></del>	<del> </del>

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
<u> </u>		<u> </u>	<u> </u>	sequence		nucleotide insertion  VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
						DTPNACATFNFLCHEGRVTGAALIPPPGGTSL TSLGQAAQ
590	1940	Α	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
						RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
591	1941	A	4495	1444	1116	VPHQGGLPGPIRVAPSSAGQREASQGPPGR IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT
						MEYYADTERNEIMSF\AGTWVELEAIILSKLM LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL
592	1942	A	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
						VCGGLSLLANAWGILSVGAKQKKWKPLEFL LCTLAATHMLNVAVPIATYSVVQLRRQRPDF EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
						MWMVCWPVNYRLSNAKKQAGHTVMGIWM GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
					:	GLGFGVCFLLLVGGSVAMGVICTAIALFQTL AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI
					•	DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL GPFSLADTHLSDLPYTWGDRDSGGACVM
593	1943	A	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC HFPASASQVAGTTHARHHTQLIF\AFLVENGL
594	1944	A	4507	1327	647	C KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
						ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII
				-		NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
						GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG
595	1945	A	4512	533	264	PADGSLYI FFFKMESYSVARLECSGAISAPCNLHLLGSNN
						SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	A	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
						NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS
						GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
						DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV
						FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS
						DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
						ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
						EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
507	1047		4510	526	924	KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV
597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE
598	1948	Α	4524	1	384	QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF TI GKLPPKTI SVKLMKATEREVAAAGVFR
L	L	L	L	L		TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RRREMQSQSVMLALRRGDAVWLLSHDHDG YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG ASELL
599	1949	A	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS DPPASASRVAGTTGARHHTQLIFVFLVETGFH \MLARDGLKLLTSSDPPASASQSSWDYRREPP RLANFFVFLVETGSRYVAQAGVQWLFTGAIP LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS VGSPKAKEALNMLTWRAEQEGGMQFWVSSE SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE GIPIMRWLSRQRNSLGGFASTQDTTVALKALS EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT HNRLLLQTAELADGTANGSV/SISANGFGFAI CQLNVVYNVKASGSSRRRSIQNQEAFDLDV AVKENKDDLNHVDLNVCTSFSGPGRSGMAL MEVNLLSGFMVPSEAISLSETVKKVEYDHGK LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLSSPWGRAVPC VSGGLPKPANITFLSINMKNVLQWTPPEGLQG VKVTYTVQYFIYGQKKWLNKSECRNINRTYC DLSAETSDYEHQYYAKVKAIWGTKCSKWAE SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSQCVTNHTLVLTWLEPNTLYCVHV ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	.1953	A	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS KPTSWAAIASKPAKPQPKMKTKSGPVMGGG LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		]		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ i		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
		1				SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ OPPO
605	1955	A -	4553	2	2304	
003	1933	<u> </u>	4555	_	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS
i						FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR
}	}	ł				DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS
						SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
						GPGLGALDRLRAHASAMGDEDLPGMAALOP
		[	1			HGVPGDGEGPHERGPPPASAPVGGTVTLRED
	İ					SAKRLERRARRISACLSDYSLASDSGVFEPLT
}	ļ					KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG
			1			SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL
		1				PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
		1		1		HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
]						LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
		ĺ				LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
						TTAQLQAVERELAEERAKLEYTEEEVLEMER
						KEEQAEAISERSWQADSVDSGCSNCTQTSPPY
					;	PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
						KVDKETNTEDLFLEEAASLVKERPSRRARGSP
[						FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL
		ļ				MARTSLDLELDLQASRTRQRQLNEELCALRE
		1	}			LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
						EAERQTRQTKLDYRHEQAAEKMLKKASKEI
		İ				YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL
	1	ł	1	٠,		PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
						VLLLQDSSGDYSLAHVREMACSIVDQKFPEC
[		1	1			GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
						EGDLIEVVLSASATFEDFQIRPHALFVHSYRA
ł I		ļ				PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH
				,		KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
		1				SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ
						SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
		ļ				CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
		l				SDDNDSERNSGLMDDMEEAMVQDAEMAMA
		[				ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP
		1				LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
j		}				KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE
		]				PLSEILSLEPVKTSALIPNGANPHCFEITTANV
[		ł	[			VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
	l	l	1			MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV
	l .				•	
		ļ				SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN
		 				VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ
	1057		4562		4400	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563	1	4499	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTPFFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL
607	1957	A	4563	1	4499	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTPFPNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	A	4563	1	4499	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI
607	1957	A	4563	1	4499	VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTPFPNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT

SEQ ID NO: of nucl-eotide scq-uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  HFNQIETLDPDSFQHLPKLERLFLHNNRITHL VPGTFNHLESMKRLRLDSNTLHCDCEILWLA DLLKTYAESGNAQAAAICEYPRRIQGRSVATI TPEELNCERPRITSEPQDADVTSGNTVYFTCR AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ EVTLRYFGSPARPTFVIQPQNTEVLVGESVIL ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS GGLYIQNVVQGDSGEYACSATNNIDSVHATA FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP EPAITWNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPD VSRNGDPFVATSIVEAIATVDRAINSTRTHLF DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY RTHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI GTETVTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN DSRARSGARCMFFVRSSPVCGSGMTSLLMNS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIVG AEIGHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD ENFQPIAQDHLPLHKAFFSFFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR DGDRLWYENPGVFSPAQLTQIKQTSLARIL.CD NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD LRVWQDCCEDCTRRGQFNAFSYHFRGRRSLE FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVNI PGACCPVCLQKRAEEKP FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
608	1958	A	4566	354	1135	
609	1959	A	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS
610	1960	A	4570	697	467	ECRGVISAH/CCTLCLPSSSDSASAF/RVARTT GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N LVIRPPRPPKVLGLQA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
ł	1	l	l	sequence	1	nucleotide insertion
611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
					1	WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
ŀ		J	J	]	1	LCAATAVLLSAQGGPVQSKSPRFASWDEMN
l			ļ		]	VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
			1			LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
		ĺ				LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
(			ĺ	ĺ		RIQHLQSQFGLLDHKHLDHEVAKPARRKRLP
i						EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
						QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR
		ì	1			RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
ļ	ļ	ŀ	j	J	]	EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
		1			:	VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
		l			•	LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
						GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
						WRGRYYPLQATTMLIQPMAAEAAS
612	1962	A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
		<u> </u>				GSPASASPVAGITGTRHHRTRG
613	1963	A	4584	687 _	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
]		1	}	j	j	SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
		}	1	,		HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
			4			LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	A	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
ĺ		ĺ	İ		Í	GGLFCAWVGTILLVVAMATDHWMQYRLSGS
						FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
-	1065		4500		ļ	GIFYMCAYRVHECRRLSTPR
615	1965	A	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
						ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
	]					PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
	1		ľ			ERRPLPKNARPSPWVPALADEWNTLHQEVTT
616	1966	A	4592	773	488	TRLPAGSQEPVKD DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
010	1,000	^*	1372	''3	700	SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
	İ		ļ			AGLELPTSGDPPALASQSAGITGVTTVPSGPG
617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
**'	.,,,,	~	10,75	٠,	1/0	DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
						REKQLQELQQQQEEEERQRQQRREERRQQNL
ļ						RARSREHPVVGHPDPALPPSGVNCSGCGAEL
		ŀ				HCQDAR*
618	1968	A	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
		l				SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
						MVGNETTYEDGHGSRKNITDLVEGAKKANG
1	1	ļ		<b>i</b>		VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA
	ŀ	1				MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
		1				HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV
	!					YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI
	1	l	1			AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
	-					AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
		1				CNPETFPSSNESRQCPNARCQFAFYGGESGYH
	l .	1	1			RALLGLQIFNAFMFFWLANFVLALGQVTLAG
(	[	ĺ	1			AFASYYWALRKPDDLPAFPLFSAFGRALRYH
	1					TGSLAFGALILAIVQURVILEYLDQRLKAAEN
		1				KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
	1	l			<u> </u>	IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
}	]	j	J			TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
1	I	l				APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
1		1				
		Ì	. !			VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
619	1969	A	4601	2	357	VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l	l	i	residue of		
	ļ	i			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	j	1	peptide		/-possible nucleotide deletion, \-possible
<u></u>	<b>_</b>			sequence		nucleotide insertion
1	ł					GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
1	í	ĺ	( !		ĺ	GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
L		<u>.</u>	l	L		NQHVECNEICHRLSLTRPSMEKPCKS
620	1970	Α	4606	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR
		1				KGHLEEEEEDGEEGAETLAHFCPMELRGPEP
<b>j</b>	ļ	J	]			LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
					Į.	TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
i	ļ	[	Į -			YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
	1	ì	<b>1</b> 1			EDTIRQTSLRERVAGSAGMAALTQDIRAALS
1	1	ļ				ROKLDHVWTDTHYVGLQFPDPAHPNTLHWV
1	!	•	1			·
						DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
	]	]		ļ		VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
			1			GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
		1				DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
1		1				NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
						KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
	l I		ł			NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA
			[	·		WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
J			<b>,</b>			PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
						HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
1			! 1			IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
1				!		AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
	[					DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
}						QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
			ŀ			LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
						IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
[			i			
	1		1			EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
1			1 .			DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\
						ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	Α	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
1		,				NTLVLKQQTFIESARSIGASDMTVLLRHILPGT
						GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
						EWGAMLNEARADMVIAPHVAVFPALAIFLTV
1 1						LAFNLLGDGLRDALDPKIKG
622	1972	Α	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
	_				- <del>-</del>	CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
] ]			]			RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
						SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL
						AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV
1 1						VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL
1 1						
1 1						LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
						TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV
<u> </u>	1055					AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
1						GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
1						DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ
, ,						IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
1						QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
						KTLNVMTDLKNAQERRKEKKRRKED*GAA
1 1						AAWSLRPSRPPSAAPSAAVCVAWASFOLTHG
						LKNRCFI
624	1974	A	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH
024	17/4	T.	4022	104	000	VICT TALQUINING PERMITTED TO THE TANK TO T
]						SLEENHFYSYPEEVDDDLICHICLQALLDPLD
						TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
				1		QHCKKSSILVNKLLNKLLVTCPFREHCTQVL
1						QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED
L						CLSPGVHHCSEV
625	1975	Α	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP
						PPLLIPSS*LSP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP ASASQVAGIAGTHH
627	1977	A	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP ARAYL
628	1978	A	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF SYKSFAVIIFFVDNTRFFSFGF
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ EPALHRVGVYLDRGTGNVSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TPFFCLSLLNGWDYRRPPHILANFFVLLVETG FHDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG
632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFITYRIITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH
635	1985	Α	4709	42	341	YIKQPDAKERRRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS RSTTVTVA*LMQKLNLSMNDAYYTVIMKMSS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide seq-	1	in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1				peptide sequence		nucleotide insertion
<u> </u>	<del> </del>			50400.00		ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
						MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
		•	!			LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
		ļ				HAG*AGLELLTSGDPPASASRSAGITGVSHHA
}						RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
		1			İ	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
				ł		YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL GOYTSOGGVTAWRKICPIFEGIGYASOMIVIL
]	ļ					LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
	1	}			}	WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA
	ļ					S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
	İ	1		ĺ		AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
				_		WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS
:		l				QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
Į	[	[		ĺ		MARSRLTATSASQVQAILLPQPPGTTDSCSPS PDHEQQPLSWVLPPPQKDMNPREQQVALGP
				1		QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
	]	1	1	]	1	LQLAASPYFSPSWAECPQPVPAGTHATWCLA
						RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
		1		ļ	ł	FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP ORDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
•			,	} -		FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
		ł				TWWFGVKFAAGGLGTFHALLNTAVHVVMY
•						SYYGLSALGPAYQKYLWWKKYLTSLQLVQF VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
		1				FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
					ŀ	QEIPLSEILRISSPRDFTNISQGSNPHCFEITTDT
		}	ļ			MVYFVGENNGDSSHNPVLAATGVGLDVAQS
	1	İ			İ	WEKAIRQALMPVTPQASVCTSPGQGKDHSK Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
						LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
}	1		]			AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
645	1005	<del>  </del>	1005	450	126	SYKDIWGWPCLCGVLHAYIPLLV
645	1995	A	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
1	1		1			PLLAGLVAADAVASLLIVGAVFLCARPRRSP
	L_ :	<u>L</u> .	:			AQEDGKVYINMPGRG
646	1996	A	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
	1				-	LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
1	1		1	İ	ĺ	HKQAVQCLKGPGQVARLVLERRVPRSTQQC   PSANDSMGDERTAVSLVTALPGRPSSCVSVT
						DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
			]	1		KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
	]		j	]	]	WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
						YYPAAVEVLHILLRGAPQEVTLLLCRPPPGAL PELEQEWQTPELSADKEFTRATCTDSCTSPIL
1						GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
					· .	EGTMGAKTERDLGPVP
647	1997	Α	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  IVMPTYDLTDSVLETMGRVSLDMMSVQANT GPPWESKNSTAVWRGRDSRKERLELVKLSRK HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
648	1998	A	4867	2030	837	QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK LKWAKDHDEEAKKIAKAGQEFARNNLMGD DIFCYYFQTFPRNMPIYK AGMLPAVGSADEEEDPAEEDCPELVPMETTQ SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI LTEQHSKRVAVILNEFGEGSALEKSLAVSQG GELYEEWLELRNGCLCCSVKDNGLRAIENLM
	·					QKKGKFDYILLETTGLADPGAVASMFWVDA ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI NEATRQVALADAILINKTDLVPEEDVKKLRT TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV SWKDDTERTNRLVLLGRNLDKDILKQLFIAT VTETEKQWTTHFKEDQVCT
649	1999	A	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV GQAGLELRTSGDPPASASQSAGITGVSHLA*P TSMPLLPFQRLCVYI
650	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR CPASFYLFLKYYLEAKFCA*GECAPSAGVGA GYKRGHKSCLLINCVVQI
651	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR PHMEPKASCPAAAPLMERKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA
653	2003	A	4965	2	283	FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTHAILVLNS*PQVICPPWPPKVLTLQA
654	2004	A	4968	3 .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

CEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			· ·	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1.	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		1	1			VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
		}			<b>J</b>	AIWWEQKRQWLLQTHWTLDKYGILADARLF
		<u> </u>				FGPQHRPVILRLPNRRALRLX*
658	2008	Α	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
						KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	A	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
039	2009	Α	3018	17	330	T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
				1		MAACWAVHVKTHMRPGLAVLPRLVLNSWS
	ł	}		Ì		*AIILLWPPKALGLQA
660	2010	A	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
		''	5020	-		HPGHLCLOPPGGPA*FLDYCRGCCPHPVPGST
į		l	1	ļ	ļ	AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH
		İ	1 .	1	ŧ	HRTGARWNH
661	2011	Α	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
						LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
		1				LELLGSSHPPTSASQSARITGVSHRAWPLK*F
						NLNQYQTLTMN
662	2012	A	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
						EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
	]				1	LHEACLGDHVACARTLLEAGANVNAITIDGV
						TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
	Ì					SPTHEGASKGHHECLDILISWGIDVDQEIPHSG TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
				ļ		WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	A	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK
005	2013	^	3000	731	500	ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
					1	GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
		l	1	i	i	WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
						QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
		j		j	]	MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
						TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
						YMAICNPLRYSVLMGHGVCMGLMAAAWAC
						GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	Α	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
	}					PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
						HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI
		l		1		RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
	İ	)	j			KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
		l				ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
		1			1	PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
<del>-</del>		1	1	1	1	VGPKAQLSGLQLQPCLYKRR
667	2017	A	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
	' '	1				PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
		1				RALPTTFADIENLKYLLFTRDASQPFYLGHTV
					1	IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
		l		J		TCADQSVIWKLSEDKQLAICLKYAGVHAENA
	1		1		1	EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
		1			1	CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
						GHYFNDTLVFLPPVGSEND
669	2019	Α	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
						ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
		1	1			RGCQHEAAPCPRGPGSDGLHHASAACASLPP
670	2000		15100	1	547	SPILPVLLPELGPL
670	2020	A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP

000 10	SEQ ID	Mot	CEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
-	uchcc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				•	sequence .	/=possible nucleotide deletion, \=possible
		Ì		peptide	ļ	
				sequence		nucleotide insertion
			ĺ	ļ		DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
						VGGTDQQYVSNDSGIYVSRIKENGAAALDGR
	ł			1		LQEGDKILSVNGQDLKNLLHQDAVDLFRNA
						GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
		L				FMVLVPVFALTMVAAWAFMRYRQQL
671	2021	Α	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF
		1				VLLLLLISLLCLYWKARKLSTLRSNTRKEKA
						LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT
			Ì		1	YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS
			1			NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA
		İ			<u> </u>	NHFVEVT
674	2024	A	5153	3	2953	LTEDQPFDILQKSLQEANITEQTLAEEAYLDA
	1	l	1		1	SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG
					]	QTLQPIGVTHVPVGASFASNTVGVQHGFMQH
		1				VGISVPSQHLSNSSQISGSGOIOLIGSFGNHPS
		1				MMTINNLDGSQIILKGSGQQAPSNVSGGLLV
						HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF
		l		1		QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIOM
						GQQNTYNVNNLGIQQHHVQQGISFASASSPQ
						GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG
		i	Ì	1		GSIVIHSPMGOPHAPQSQFLIPTSLSVSSNSVH
	ļ					HVQTINGQLLQTQPSQLISGQVASEHVMLNR
	ŀ	l				
					1	NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ
	ţ	1				TFAASGSPVIANHASPQLVGGQMPLQQASPT
	1.					VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR
			ļ	ļ		FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL
						TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT
				•		GTQQFFCQAQKKCLNQTSPISAPKTTDGLR
		1				QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ
					1	QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ
	ł	l		ł		LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD
			ļ			AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV
						ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE
						MVMIDRMFNQEERASLSRDKRLALVDPEGFQ
	l	ŀ				ADFCCSFKLDKAAHETQFGRSDQHGSKASSS
		1			1	LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP
	}	Į.	}	1	1	NHIVVSAEGNISKKTECLGRALKFDKVGLVQ
	1	ļ		1	]	YQSTSEEKASRREPLKASQCSPGPEGHRKTSS
		1			}	RSDHGTESKLSSILADSHLEMTCNNSFQDKSL
	1	1	1		1	RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET
	l			1		TFKNILELKKAGRQPQSDPTVSGSVELDFPNF
				1	1	SPMASQENCLEKFIPDHSEGVVETDSILEAAV
	]	}	]	J	J	NSILEC
675	2025	A	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR
	-	1	1			LYDEVQEVVYFPAVVHDNLGERLKCTYIEID
	ŀ			[	1	QVPETYAVVLSRPAWLWGAEMGANEHGVCI
		1	1			GNEAVWGREEVCDEEALLGMDLVRLGLERA
				1	1	DTAEKALNVIVDLLEKYGQGGNCTEGRMVF
		1		Į.	Ī	SYHNSFLIADRNEAWILETAGKYWAAEKVQE
	1	1	(		1	GVRNISNQLSITTKIAREHPDMRNYAKRKGW
	1			1	1	WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE
				1		GYKLLNKHKGNITFETMMEILRDKPSGINME
	1			1		
		Į	}	1		GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER
			ļ ,	<b>l</b>	1	SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI
			1 .	l .	I .	I HEKPUKKHPL YUKHUUALEVVNNNEEKAKI
			1			A DAN ON PREI PAR SES SES SES
						MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

			1 5 5 5	<del></del>		(A. Alesian C. Contribution
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	in	nucleotide	location	
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Į.				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l	}	Ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ľ	ì	peptide	1	/=possible nucleotide deletion, \=possible
		<u></u>		sequence		nucleotide insertion
676	2026	A	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG
1						FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG
	ł	ļ	1			FTLLARMVSIS*PHDPPASASQSAGITGVSHRA
						RPT
677	2027	Α	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC
		i			ļ	KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF
ł	1					SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR
i					1	LNKRSFFMISPTDQQVHCWAWLKKHMPKDS
ł	i	ľ	ľ			NLLLEDVTWKYTALNLIGPRAVDVLSELSYA
				l		PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT
			L			GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	Α	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ
						GRIAKMPVKWIAIESLADRVYTSKSDVWAFG
l	İ	1	i	1	l ·	VTMWEIATRGMTPYPGVQNHEMYDYLLHG
						HRLKQPEDCLDELCKI**SPQSP
679	2029	Α	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE
		ļ		Į.		KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH
	!					VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI
1	1	į.				EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS
]			J			AFDHFASVHSVSAEGTVVSNLSS
680	2030	Α	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL
				1		LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH
		İ				RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL
			1		1	FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG
	J	j		j	)	LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI
						KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF
ŀ	1	1		}		DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F
]	J	J	j			SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC
Į.		į				WPGWSRTPDLS
683	2033	Α	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF
		1		l		MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG
1			1			YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK
]	J	j	}	ļ	1	QSESAI
684	2034	Ā	5220	1	194	NLMKEMQNLNSENHKTWEEYKDTK*IMSYF
		1	1			YG*ALNVIKMAVLPKLMYRFSATLVKJPQHL
						TDS
685	2035	A	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED
				1		QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS
		{		F	1	AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR
				1		KHSRPIVTVWERELRKAKPNRKLTFLYLAND
			1	Ī	1	VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD
			1	1		ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA
		**	1	1 -	1	NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI
		1		i		PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR
			1	I	1	EVDKDRVKQMKARQNMRLSNTGEYESQRFR
				}		ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP
1 000	2030	_ ^	1 2247	1 *	1407	SGDSDLATALHRLSLRRQNYLSEKQFFAEEW
					1	QRKIQVLADQKEGVSGCVTPTESLASLCTTQS
1		l			1	EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY
		ŀ	1		1	HWQQLAQPNLGTILDPRPGVITKGFTQLPGD
1		1	,			AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS
J	1	1	Į.	}	}	
	1	1				
				<u> </u>		KPVTGIFLPPITSAGGPVTVATANPGKCLSCT
						NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ ID	SEQ ID	Met	L CCO	1 Descriptor	Description of the second	
NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l	l	1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j		1		peptide	]1	/=possible nucleotide deletion, \=possible
	_			sequence		nucleotide insertion
	1					FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
1	ł	1			1	PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
		ł				FLASRPAETFLQEMYGLRPSRNPPDVGQLKM
						NLVDRLKRLGIARVVKNPGAQENGRCQEAEJ
		1	ĺ			GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM
		<u> </u>			}	GSFAAPVCTSSPKMGVLKED
689	2039	A	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS
		1				GAPAGARGGPAKANSNPFEVKVNRQKFQILG
	İ	1				RKTRHDVGLPGVSRARALRKRTQTLLKEYKE
	1	ł				RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
	1	J	1	j	ļ	LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE
						KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
			1	1		GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
}	1	ł	1	1	}	QEKRERQAQREDALELTEKLDQDWKEIQTLL
					1	SHKTPKSENRDKKEKPKPDAYDMMVRELGF
	1	{			-	EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
	1	ł	1		1	RLRRMLGKDEDENVKKPKHMSADDLNDGFV
	1	ł	1	1		LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA
						SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
		ĺ	İ	[		NVESEEENEKPAKEOROTPGKGLISGKERAG
	ł	l		<u> </u>		KATRDELPYTFAAPESYEELRSLLLGRSMEEO
						LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
						YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP
		1				ESASDAIKFVLRDAMHEMEEMIETKGRAALP
		Ì	ł			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
	ļ	]		ļ ·		SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
	-	ŀ				QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
			1			ALGKNSELLVVSAREDVATWOOSSLSLRWA
	1	1				SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
	,		J			YGSLPSFHAIMGPLRALLTDHLADCSHPOELO
	<b>[</b>					ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL
						FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK
	1					REFKGAVREIRKDNQFLARMQLSEIMERDAE
<u></u>	2242					RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	A	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
	} ,		1 1			ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
						FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
601	204:		- <u></u> -			SFVK
691	2041	Α	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
602	2040					EVLSSFFFFFLKFSYKPQNIV
692	2042	A	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
					ļ	ERVLTFLPAKALLRVACVCRLWRECVRRVLR
			[		i	THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
	1 1					ENVRILPHTVLYMADSETFISLEECRGHKRAR
	į l		] ]	İ		KRTSMETALALEKLFPKQCQVLGIVTPGIVVT
				l		PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF
	i 1			ł		HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
				ł		FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
						QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
					j	QSATVLLNEDVSDEKTAEAAMORLKAANIPE
	[					HNTIGFMFACVGRGFQYYRAKGNVEADAFR
	}				j	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
602	I		L			EVKDDDLFHSYTTIMALIHLGSSK
693	2043	A	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
606						ACFPTNIVTLCHSIA
694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
	1 1		-	1	į.	KTRAMRRRLNMHEENLKTKKQHRKERLYPL
	ļI	I				
695	2045	A	5315	125	1596	RKYAAKA ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

NOO of NOO of Incident Continued and the Common of Noo of Incident Continued and Incident Continued Contin	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide   seq							D=Aspartic Acid. E=Glutamic Acid.
Sequence			1100				F=Phenylalanine G=Glycine H=Histidine.
1949	-	1					
uence    914   ng to first amino acid acid residue of peptide	1		1				M=Methionine N=Asparagine P=Proline
amino acid residue of peptide residue of peptide residue of peptide sequence   T=Threeonine, V=Valine, W=Typtophan, Y=Typtophan, Y=Stop codon, Y=possible nucleotide deletion, V=possible nucleotide deletion nucleotide deletion nucleotide deletion nucleo		uchee		-			
Persidue of peptide   Sequence   Y=Tyrosine, X=Unknown, **Siop peodon, /*possible nucleotide (insertion   Persidue)   Persidue nucleotide (insertion   Persidue)   Persidue	dence			714	_		
peptide	ł						
	]				1	sequence	
			İ				
SSLFESLISPIKTETFKEFWEGKPLLIGRIDPPA	ļ				sequence		
	1						
					Ï		
ATIOFHQPQRPKDELWRIGERECTYGSLVGS   NYYTITPAGSGGLPHYPTYDDVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYPTODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDODVEVH_QLGEBCH   WRL YJPITVPI AGROGLPHYDTAKEDVELRTG   IPRQLLQVESTIVATRRLSGFLRTLADLEG   TRELISDMKKDFHMRLPYSAGOGAELSTP   GGKLPRLDSVVRLQFKDHIVLTVLPDQDQDG   ETQEKMY1YTHSLKNSRETHMKGNEEFEFH   GLRPFLSHLJGLWCHANGREEFH   GLRPFLSSSPELAKRSWFONFISLOKE   GLPFLJCHGWCHANGREEFEFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GLRPFLJCHGWCHANGRAUGH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASSYPCIFCEBFH   GRASS							
NYTITPAGSGGLPHTYDDVEVFILQLEGEKH   WRLYPHTYPLARFSVSAEBEIGRYHEFML   WRLYPHTYPLARFSVSAEBEIGRYHEFML   KPGDLLYFPRGTIBIQADTPAGLABSTHVTIST   YQNNSWGDFLLDTISGLPTDTAKEDVELRTG   IPRQLLLQVESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKRDFIMRLPPYSAGDGAELSTP   GKLPFLDSVYRLQPKDBIVLT-UPDQDQSD   ETQEKMYLTYHSLKNSRETHMMGNEETFET   GLRPFLSHLDALKQIWNSPAISVSKULKLTTDF   GLRPFLSHLDALKQIWNSPAISVSKULKLTTDF   GLRPFLSHLDALKQIWNSPAISVSKULKLTTDF   GLRPFLSHLDALKQIWNSPAISVSKULKLTTDF   GLRPFLSHLDALKQIWNSPAISVSKULKLTTDF   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQV   EKESLVLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTLSLWTECLIQV   EKESLVLSLWTLSLWTLSLWTECLIQU   EKESLVLSLWTLSLWTECLIQU   EKESLVLSLWTLSLWTECLIQU   EKESLVLSLWTLSLWTSLWTLSLWTSLWTLSLWTSLWTSLUDIKTVECLIST   EKESLVLSTAT						· ·	
WRLYHPTYPLARRYSVEAERIGR7VHERM   RFGDLTYPRGTHIQADTRAGLARSTHVTIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNNSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNDSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNDSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNDSWGDFLIDTISGLVPDTAKEDVELTRIST   YQNDSWGDFLIDTISGLVPDTAKEDVELTARDEG   YEELLSSDWGTKLARDEGPDQNSDLSTENGEFTH   GREFILDSWGTALDSWGDLSTRIST   GREFILDSWGTALDSWGDLSTRIST   YEELLSSDWGTTHLTARSSGGTET   SQLQUBSKLJSCHGTAKENSKALKEEMEGVVKE   SQLQUBSKLJSCHGTHYDTYSSTILVENTLGENKKIKEEMEGVVKE   SQLQUBSKLJSCHGTHYDYSSTILVENTLGENKKIKEEMEGVVKE   SQLQUBSKLJSCHGTHYDYSSTILVENTLGENKKIKEEMEGVVKE   SQLQUBSKLJSCHGTHYDYSSTILVENTLGENKKIKEEMEGVVKE   SQLQUBSKLJSCHGTHYDYSSTILVENTLGENKKIKEEMEGVVKE   SQLQUBSKLJSCHGTHYDYSSTILVENTLGELTLOSTAGSC   SQLQUBSKLJSCHGTHYDYSSTILVENTLGELTLOSTAGSC   SQLQUBSKLJSCHGTHYDGLGSLGGP   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSLGGF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGTHYDGLGSKFF   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQUBSKLJSCHGT   SQLQU							
YQNNSWGDFILDTISGLVFDTAKEDVELRTG   PROLLLQVESTVATRALSGILAGENTENT ADRILEG   TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP   GGKLPLSDWKKDFIMHRLPPYSAGDGAELSTP   GGKLPLSLLDSVVRLQFKCHIVLTVLPQDQDSD   ETQEKMVTYYSLKINSRETHMMONEETEFH   GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE   EXESLVLSLWTECLIQVV	ļ,	[			}		
PRQLLLQVESTTVATRRLSGELRTLADRLEG   TEELLSSDMKKDFMMRLPPYSAGDGAELSTP   GGRLPRLDSVVRLQFKDHIVLTVLPDQDQSD   FTOEKMYVTJUSLKNSRETHRMGHEETEFH   GLRFPLSHLDALKQIWASPAISVKDLKLTTDE   ERESLVLSLWTECLIQVV							
TKELLSSDMKKDFIMIRLPPYSAGGGALSLTP	1						YQNNSWGDFLLDTISGLVFDTAKEDVELRTG
GGKLPRLDSVVRLQFKDHIVLTVLPQDQSD							
GLŘFPLSHLDALKQIWNSPAISVKDLKLÍTDE						ļ	GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD
	l	l			i		
696         2046         A         5318         1476         742         LMKXYLEAAELGEISDIHTKLIRLSSSQGTIET SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN VEKLQVLLNCMTEITYQFKKDKAERRLAYN EGQIHKFDKQKLYYHATKAMTHFTDECVK YEAFLNKSEEWIRKHLRRQLLSLTHQKMTTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHLLBSGGSLTHOGGLRNVDCL           697         2047         A         5320         244         478         LDYNFFLFENTFGLVSQAGVWHDLGSLQPF PGFKQFSCLSLPSSWDYRHLPHLANFSREG VSPSWPGWSRTPDFR           698         2048         A         5324         266         714         LPIRKSLRSVRSGFTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQHRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVDKQDKD           699         2049         A         5334         699         277         RPIGHLVCISSSAGLSGVNGLADYCASKFAA PGFAESVFVETYVQKQKGIKTTVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALIQEKM YLYMPKLYFMMFLKSFLPLKTGLLLDYLGI LHAMDGFADQKK           700         2050         A         5344         3         614         PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLYLKDKPLSSIKSVLSQTSYTTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTHTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTHLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLLSTRUTYTTLISGPSRFKRV VETIQAGLISTRUTYTTLISGPSRFKRV VETIQAGLISTRUTYTHE VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETICATION VETIC		ļ					GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
SLQDIDSRLSPGSLADAWAHQEGTHPKDRN	1		ļ				EKESLVLSLWTECLIQVV
SLQDIDSRLSPGSLADAWAHQEGTHPKDRN	696	2046	Α	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
VEKLQVILINCMTEIYYQFKKDKABERRLAYN   EEQHKFDKQKLYYHAKMHFIFDECVKK   YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI   EEEVSKYQEYTINELQETLIQKMKTASSGKHT   TIPTYPSSNTILVEMILGMKKI KEEMGVVKE   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGGSLTMDGGLRNVPCL   LAENNHILESGESTART   PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG   VSPSWPGWSRTPDFR   VSPSWPGWSRTPDFR   VSPSWPGWSRTPDFR   VSPSWPGWSRTPDFR   VSPSWPGWSRTPDFR   VSGLIMAFQKYSGETVQERKQKDRKALHEL   KLEEWKGRLQVTEHLPEKIESSLQEEDEPENDA   KKIEALLNLPRNPSVUDKQDKD   KKIEALLNLPRNPSVUDKQDKD   KKIEALLNLPRNPSVUDKQDKD   KKIEALLNLPRNPSVUDKQDKD   KKIEALLNLPRNSVUDKQDKD   FPHGHLVCISSSAGLSGVNGLADYCASKFAA   FGAESVFVETFVQKQKGKTTIVQPFPKKTGAL   PHHGHLVGISSSAGLSGVNGLADYCASKFAA   FGFAESVFVETFVQKQKGKTTIVQPFPKKTGAL   LAHNDGFADQKK   TYTAFEMSSLTPESSPELAKRSWFGNFISLDKEE   QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS   QTSFRAEYKASGGFSVPQKFVRPKVPRQVIDISSEG   PESPSRRDGSGGGGTSVTVTTILISGFSRFFKRV   VETIQAQLLSTHDQPSVQALADEKNGAQTRP   AGAPPRSLQPPPGRPDFDELSSPRRGPKDKK   LATNGTPL   GSPSVPCIFCEEHFPVABQDKLLKHMIEHKIV   LADVKLVADFQKYLLVARKRTEOPITIDECS   RINSTAPFEQENYFLLCDVLPEDRILREELQ   KQRLREILEQQQQERNDTINTHGVCMFCNEEF   LGNSVVLINIMAREHAFNIGLPDNIVNCEFL   CTLQKLDNICALPSLEFSGGTTTLE   GSPSVPCIFCEEHFPVABQDKLLKHMIEHKIV   LADVKLDVADFQKYLLVNFRKQVCMFCNEEF   LGNSVVLINIMAREHAFNIGLPDNIVNCEFL   CTLQKLDNICALPSLEGSGGTTILE   GSPSVPCIFCEEHFPVABQDKLLKHMIEHKIV   LADVKLDVADFQKYLLVNFRKQVTHCPCNEEF   LGNSVVLINIMAREHAFNIGLPDNIVNCEFL   CTLQKLDNICALPSLEGSGGTTTLE   GSPSVPCIFCEEHFPVABQDKLKKSHEG   NFYQQVKLNVFIKRGVHQCCYGCHVKFKS   KADLRTHMEETKHTSILPDRKTWDQLEYYFP   TYENDTILLWTLSDSSEDLTAQEQNENVPIISE   DTSKLYALKQSSLBOLLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLAQEQNENVPIISE   DTSKLYALKQSSLBOLLA	1			•			
BEQHKPDKQKLYYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILLESGGIVHTDGGLRNYDCL     697   2047   A   5320   244   478				İ	ļ		
	i			Í	ĺ	ĺ	
BEEVSKYQEYTNELQETLPQKMFTASSGIKHT   MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE   LAENNHILESGGSLTMDGGLRNVDCL		ŀ					
MTPIYPSSNTLVEMTLGMKKLKEEMEGVYKE	ļ						
LAENNHILESGGSLTMDGGLRNVDCL	l	1					
697         2047         A         5320         244         478         LDYNFFLFEMTFGLVSQAGVQWhDLGSLQPP PPGKQFSCLSLPSSWDYWRHLPPHLANFSREG VSPSWPGWSRTDFR           698         2048         A         5324         266         714         LPIRKSLRSVRSGFTTSQSFTTNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD           699         2049         A         5334         699         277         RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLIPLBPKYAVEKIVEALQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFAQKK           700         2050         A         5344         3         614         PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDSGGGITSVTFTTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPRSLQPPGRDPSLSSSPRRGPKDKK LLATNGTEL           701         2051         A         5346         3         1383         HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLIEPLSLPESPGGTTTLE GSSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILVWRKRFTEQPITDFCSV IRINSTAPFEEQQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRSUFNKREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRRTWDQLEYYFF TYPENDTILWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLI	1			ĺ		ł	
PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VSPSWPGWSRTPDFS VGGLLMAFQKYSGEFTVQERKQKDRKALHEL KLEWKGRLYOFEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGYNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVFQVDISSEGG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPSLQPPFGRPDFELSSPRGFPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGGDSKDCLLEPLSLPESSPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILLTQHTDFCSV RINTSTAPFEEQENYFLLCDVFPDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTIQKKLDDNI-LYVTKPRTEELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTIQKKLDDNI-CLYCEKTFROKNTLKDHMR KKQHKNPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHFASAVCL FCEKQAETTEKLYVHMEDAHEFDLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTUNDQLEYYFF TYPKNDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	607	2047		5320	244	178	
VSPSWPGWSRTPDFR	697	2047	1 ^	3320	244	770	
698 2048 A 5324 266 714 LPIRKSLRSVRSGPFTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLMAFQKYSGETVQERGVEDKRÄLHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKQNKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVORLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGFSVFQKPVRFQVDISSSEG PESPRDGSGGGGIYSVTFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPFGRPDPELSSSPRRGPKDK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESSFGSTTILE GSPSVPCIFCEEHPPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKFTEQPITDFCSV RRNSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLDDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTILKDHMR KKQHRKINPKNREYDRFVVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETELLYVHMEDAHEFDLIKKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1		l	ļ	ł		
LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALLQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYXASGGPSVFQKPVRFQVDISSSEG PEFSPRRDSGGGGIVSYTFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGGNSVILNHMAREHAFNIGLPDNIVNCNEFL LGRSVILNHQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFF TYENDTLLUWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	600	2048	-	5224	266	714	
VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALIQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGYGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHIPPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRISVINHIMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	098	2040	^	3324	200	/14	
KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLYLKDKPLSSIKADIVHAFLSIPSLSHSVLS. QTSFRAEYKASGGPSVFQKPVFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFIL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILLPDRKTWDQLEYYFT TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1	1				
KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPSLQPPFGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNEYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1				
699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKILLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTITLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEHIFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEEL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWILSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSILPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWILSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		2010		5004	-	000	
FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEFSPRRDGSGGGIYSVTIFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSSLG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	699	2049	A	5334	699	211	1
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHIFVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL				ļ.		1	
LHAMDGFADQKK		ļ					FEGUT GCPSLLPILEPKYAVEKIVEAILQEKM
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTIFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYIL YWRKRFTEQPITDFCSV IRINSTAPFEEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHIMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTIFTLISGPSRRFKRV VETIQAQLLSTHIDQPSVQALADEKIGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		l	<u> </u>				
QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTIQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	700	2050	Α	5344	3	614	
PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1					
VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1				PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
LLATNGTPL     701   2051   A   5346   3   1383   HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1				VETIQAQLLSTHDQPSVQALADEKNGAQTRP
701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1						AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					· .	1	
STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	701	2051	A	5346	3	1383	
GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1		ļ		
IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL				1	i	1	
IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	}				1	1	
KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			ļ		1	i	
LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1	1		Į	1	
CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		}	1	l	1	}	
KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1		1	ļ	1	1	
VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1		Į	!	
FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1			1	1	
LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1				
KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1				
TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1		1			1	
DTSKLYALKQSSILNQLLL		ļ	]		}	1	
	1	İ	1				
702   2052   A   5356   2502   1540   MAAATRGCRPWGSLLGLLGLVSAAAAAWD		<u> </u>				L	
	702	2052	Α	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			}			LASLRCTLGAFCECDFRPDLPGLECDLAQHL
		ļ				AGQHLAKALVVKALKAFVRDPAPTKPLVLSL
						HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA
			<u> </u>			CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV
ļ	ł	}	}	'		VYGTNYRKAIFIFISNTGGEQINQVALEAWRS
						RRDREEILLQELEPVISRAVLDNPHHGFSNSGI
			ļ	•		MEERLLDAVVPFLPLQRHHVRHCVLNELAQL
l			l			GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK
						TVASRIAFFL
703	2053	Α	5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL
		l				EERLEFWMEKYDKDTEMKQNELNALKATKA
ļ		ļ				SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK
	1		1			KVKQDLLELKSVIKLQAWWRGTMIRREIGGF
704	2054	A	5381	1	1003	KM FRGRAVKMAAVVEVEVGGGAAGERELDEV
,,,,,	2034	^	7581	<b>*</b>	1003	DMSDLSPEEQWRVEHARMHAKHRGHEAMH
Ì						AEMVLILIATLVVAQLLLVQWKQRHPRSYN
			}			MVTLFOMWVVPLYFTVKLHWWRFLVIWILF
1	ł	ł				SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY
						KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA
1	Į		[	:		MDFGISLLFYGLYYGVLERDFAEMCADYMA
	ĺ					STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV
			İ			SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK
1			ł			QTCPYCKEKVDLKRMFSNPWERPHVMYGQL
705	2055	A	5396	3	675	LDWLRYLVAWQPVIIGVVQGINYILGLE IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP
""	2033	^^	3370	~	073	GNKRPLSALYRLESKEPFLSVGGYVFDYDYY
					•	RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA
1			[			VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS
						DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK
						AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP
						AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ
706	2056	A	5410	2	98	K GRVGLNLEGRGCSEPKWRHCTPTWATEODSI
/00	2030	^	7410		90	S S
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT
'				Ĭ		VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP
	1	)	]			EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	Ā	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS
			1			TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS
	<u> </u>	<u> </u>				GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG
						SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ
			1			VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK
710	2060	A	5442	1073	559	GKWFCPRCVQEKRKKK
'"	2000	l ''		10/3	339	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY
1	1	1	ł			EESDSDESWTTESAISSEAILSSMCMNGGEEK
1		l	1			PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI
		1	l			RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV
	<u> </u>	<u></u>	L			SAEIIRKMQQ
711.	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP
						DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL
1	1	1	1			NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD
712	2062	À	5400	01	740	LYRDIPELQGF
/12	2062	^	5499	91	749 ·	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS
1	l	l	l			DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF
	<u> </u>	Ц	<u> </u>	ــــــــــــــــــــــــــــــــــــــ	L	TOT TOTAL DOLL OF YOR VOILE AAPT

SEQ ID	SEQ ID	Met	LCEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						KAPELLQGQSEDEQPDASQMHVYSLGMTLY
[		[				WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH
ŀ			Į			RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR
712	2063	<del> </del>	5506	22	478	RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
713	2003	Α	5506	22	4/8	VEELILVSRLDPHILHTPMYFFLAHLSFLDLSFT TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL
						GGVECLLLAVMAYDRCVAICKPLHYMVIMN
						PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR
						CGHHEVDHFLCEMPALIRMACISTV
714	2064	A	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR
114	2001	^	3314	23	220	RLTSSVSCALDEAAAALTRMRAESTANAGQS
		l	İ	ĺ		DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM
·				] -		GRTALFHHSGGSSGYESLRRDSEATGSASSAP
						DSMSESGAASPGARTRSLKSPKKRATGLQRR
		[		[		RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG
				ł		SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG
		į.		ļ		LACVSTRLRLAERRQQRLREVQAKHKHLCEE
						LAETQGRLMLEPGRWLEQFEVDPELEPESAE
		[	[	ĺ	[	YLAALERATAALEQCVNLCKAHVMMVTCFD
	1	<u></u>				ISVAASAAIPGPQEVDV
716	2066	A	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD
		1		1		SSSSERTMDLVLEMCNTNSIHWCGISGRQLG
				ŀ		KLHPSSSI CLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG
/1/	2007	^	3331	[ ]	1 400	RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS
						VNIREELKLANVVFFPRCLLVQRCGGNCGCG
		1			j	TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR
						RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF
						TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS
						KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM
	i	ł	}	ì		LATNSTRGLNEDELMAHGQEKDSSSESEDSC
				ł		PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL
L	0.050	<u> </u>	1			FPLAFGLFNIVAAERC
720	2070	A	5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP
		1	1	ĺ	[	LLLRCKLLPSALQKMALGMFFGFTSVIVAGV
		1				LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG
						AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH
					1	CPKDFGNINNCRMDLYFFLLAGIQAVTALLF
	1	ĺ	[	ĺ		VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA
, 1	-3/1	••	5552	1	330	KLHETALHHAAKVKNVDLIEMLIEFGGNIYA
		ļ		1		RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT
						LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID
			1	}	ł	YLSYN
722	2072	A	5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF
	1					TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES
	1		1			LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD
	ł	ł	1	ł	1	VALPALGAVLRGAVRVEKNQELCHLSTIDW
						GLLQPAPGANHIVGNKLGEECADVCPGVLGA
		1			-	AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP
	[			1		HGMACTARGECCHTECLGGCSQPEDPRACV
		1	}			ACRHLYFQGACLWACPPGTYQYESWRCVTA
ļ	1	1	1			ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT
1	1		l	<u></u>	l	RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

			1 050	<b></b>		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Boquoneo	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				sequence	ļ	
						QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL
						GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
						AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
					ļ	PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN
		1				KAEINPRTNGDRAACOTRTLRFVSNVTEADRI
						LLRWERYEPLEARDLLSFIVYYKESPFQNATE
					}	HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
	<b>j</b> .	}	l .	ļ	ļ	ASLKPWTQYAVFVRAITLTTEEDSPHQGAOS
		1	1			PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
						KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
	· .		ŀ			YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD
		[			<b>l</b> .	CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
			1		1	NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL
	1	I	1			RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY
	ľ					RIDIHACNHAAHTVGCSAATFVFARTMPHRE
				l .		ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
	i	ì	l		1	ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
	1	l	Ì			HLALLPPGNYSARVRATSLAGNGSWTDSVAF
		ļ		-		YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
	1	]		1	į	LGFFYGKKRNRTLYASVNPEYFSASDMYVPD
	·	l				
		ļ			]	EWEVPREQISIIRELGQGSFGMVYEGLARGLE
	1					AGEESTPVALKTVNELASPRECIEFLKEASVM
		ŀ			]	KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
				1		GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
	1					AGEIADGMAYLAANKFVHRDLAARNCMVSQ
			1		[	DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
		l				VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV
	Į	l				TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
i						PLOLOELMSRCWOPNPRLRPSFTHILDSIQEEL
					ł	RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
1			İ			TPRDCSPQNGGPGH
723	2073	<b>-</b>	5672	ļ., ——	216	
123	2073	A	36/2	1	216	LAWIDNILPEKEKKETDKKRKKKGAHEDCD
	İ	1				EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR
		<u> </u>				SSSWSYSL
724	2074	Α	5704	4235	940	ARGRESRPVWAASWGGRGRPAARREPRGLA
ĺ	1					ATMGFELDRFDGDVDPDLKCALCHKVLEDP
	1		}	I	1	LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
1		1		1	1	LSAKELNHVLPLKRLILKLDIKCAYATRGCGR
		1	[	1	1	VVKLQQLPEHLERCDFAPARCRHAGCGQVLL
1		1		[	1	RRDVEAHMRDACDARPVGRCOEGCGLPLTH
		1		1	1	GEQRAGGHCCARALRAHNGALQARLGALHK
					1	
	1			1	1	ALKKEALRAGKREKSLVAQLAAAQLELQMT
1	1	l	1	1	ł	ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE
l		ł	ŀ	1		ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG
1				1	]	SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN
1		1	Į	I	1	GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT
		1	]	I		PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
				1		KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
			]	1	1	GDIHQEMDREELELEEVDLYRMNSQDKLGLT
1		1	1		<b>!</b>	VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG
1	1		1	1	1	DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
}		J	J	1	J	
1		1		1	1	ARAELQLDEGWMDDDRNDFLDDLHMDMLE
		1		ļ	į.	EQHHQAMQFTASVLQQKKHDEDGGTTDTAT
1			1	1	I	ILSNQHEKDSGVGRTDESTRNDESSEQENNG
		1	1			DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
1	1	4				•
		1				NESFISADCTDADYLGIPVDECERFRELLELK
						CQVKSATPYGLYYPSGPLDAGKSDPESVDKE
						CQVKSATPYGLYYPSGPLDAGKSDPESVDKE LELLNEELRSIELECLSIVRAHKMQQLKEQYR
						CQVKSATPYGLYYPSGPLDAGKSDPESVDKE

SSQ ID Not of much colded sequence peptide solid sequence (sequence peptide solid sequence peptide solid sequence peptide solid sequence peptide solid sequence (sequence peptide solid sequence peptide solid sequence peptide solid sequence peptide sequence (sequence peptide sequ		T				- · · · · · · · · · · · · · · · · · · ·	(1 A)
nucleotide sequence   1944   1958   1944   1958   1944   1945   1944   1945   1944   1945   1944   1945   1944   1945   1	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			hod			I	D=Aspartic Acid, E=Glutamic Acid,
	nucl-	peptide			l.	1.00	
1	eotide	seq-		USSN	location	corresponding	
amino acid residue of peptide residue of peptide residue of peptide sequence   T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, Selo codon, deposible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide deletion, \=\text{prossible nucleotide insertion} } \]    ABGISCPSSIGAVGTTEAYOFASKULISITE	seq-	uence	l	09/496	correspondi	to last amino	
amino acid residue of peptide residue of peptide sequence   T=Threeonine, V=Valine, W=Tryptophan, Y=Tyrosine, Section condition, Possible nucleotide deletion, V=possible nucleotide deletio	uence	1	j	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Pepside of peptide sequence		1				of peptide	
peptide	i	ľ	i	ł	1	1 1 1	
				i		Soquonoo	
AAEGISCYSSEGAVOTTEAYGPASKNILSITE   DPEVGTPYSPSKLKLDPNOPLESKERASDO						•	
DPEVGTPTYSSIXELDPNOPLESKERRASDG   SRSPTPSCKLGSAVLPSYAOSOMSLVSMCK   DLSSPTSPERMEWKVKKRSDGTRYTKRPVB    DRLIRERALKREERSGMTTDDDAVSEMKM    GRYVSKEERKQHLVXAKERSGRTSKTKRPVB    DRLIRERALKREERSGMTTDDAVSEMKM    GRYVSKEERKQHLVXAKERSGRTSKTKRPVB    DRLIRERALKREERSGMTTDDAVSEMKM    GRYVSKEERKQHLVXAKEQRRREERSEMGSR    LDCLKEQQAADDREMMILEISHKKMMKKR    NKRIPDNWAMTDGLLTHGTKSPPGTRVTNSF    LSVTTV    LSVTTV    LCGKNGGGAFLIPYFLTLIFAGVPLFLLEGSLG    QYTSIGGIGWKLAPPKGVGLAAAULSFW    LNYTYNISWAIYYLYNSFTTLPWKQCDNP    WNTDRCFSNYSMVNTYMSMILEFRGW    LNYTYNISWAIYYLYNSFTTLPWKQCDNP    WNTDRCFSNYSMVNTYMSTATYPMILLIFFRGV    CINKGTSMFAGFVFSIVGFMAHVTKRSIADV    ARSOFICHALLSTAVALUSFW    LNMYMISWAIYYLYNSFTTLPWKQCDNP    WNTDRCFSNYSMVNTYMSITULFFRGV    CINKGTSMFAGFVFSIVGFMAHVTKRSIADV    ARSOFICHALLSTAVALUSFY    ARSOFICHALLSTAVALUSFW    ALMGIDSQFCTVEGFTALVDEYPRLLENRR    ELFAAVCISTY-IGLSINTGGGIVYKLAPTYMILLFFRGV    TPGVGGLGAFLAYPEAVTQIPSPFWAULFFSW    ANGENIEWSLCWSFFTPIVAGYVFSDIVC    CINKGTSMFAGFVFSIVGFMAHVTKRSIADV    ARSOFICIWSLCWSFFTPIVAGYVFSDIVC    CINKGTSMFAGGVFSTFTPIVAGFVFSDIV    ARSOFICIWSLCWSFFTPIVAGFVFSDIV    ARSOFICIWSLCWSFFTPIVAGFVFSDIV    ARSOFICIWSLCWSFFTPIVAGFVFSDIV    ARSOFICIWSLCWSFFTPIVAGFVFSDIV    ARSOFICIWSLCWSFFTPIVAGFVFSDIV    RPENGEPOPOAGSTSTREAY    TPGVMAYMFLTLKOSIS, ORIONOPSEDIV    RPENGEPOPOAGSTSTREAY    TPGVMAYMFLTLKOSIS, ORIONOPSEDIV    ARSOFICIWSLCWSFFTPIVAGFVGFWAHVAINAVQNCAII    CVNDOPCEMTOFSRPDVMQRFCTCD    TPGVAGAVGFTPIPLAGFTGANPYVERGHVAL   AWFEKMRCYLQLLFYTLHFKTRL   AWFEKMRCYLQLLFYTLHFKTRL   AWFEKMRCYLQLLFYTLHFKTRL   AWFEKMRCYLQLLFYTPLHFKTRL   AWFEKMRCYLQLLFYTPOPPDLCGRITTC   YMASKNSSGFTCTRARELAQOGSHHISINID    AVAVAVMGIFSLTYGSSFPATATIVOGLVUQLLRSWGWHG   LVLGSANVDESLIGTVTAPYTSSMPTIGGSPENIAL   ALQVVQARIRMVLAYLFAQLSILWSFRTYTHNFALDA   AGNIPALSTRARAPQFGAAGAGKKLADFGFP    PASRRLRAPGSRPTLAPTTRAAQPAHARMA    PANGAVMGIFSLTYGSSFPATATIVOTHNINCLUCHANSKORITTARELAQOGSHHISINID    AVAVAVMGIFSLTYGSSFPATATIVOTHANDALDAAGTVTYLAVNNINCLUCHANDAAAGTVTYLAVNINCLUCHANDAAAGTVTYLAVNINCLUCHANDAAAGTVTYNNKLI    CLILICARAGAGALGCGRIFTSTTTNNFALDAAAGTTAAATTTAAATTTAAAQPA				1	sequence		
SRSFTFSQKLGSAYLPSYHHSPYKHAHIPAMIC   OHYOSYMQLOQKSAYEYAQSOMSLVSMCK	ŀ		İ			<b>\</b>	1
		ł		1			DPEVGTPTYSPSLKELDPNQPLESKERRASDG
DISSPTMSERMENT   DRILITERALKINESSOMTIPDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUSKEMMILGSIKKMMKKR   NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF   LSVITTV		ŀ				1	SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
DISSPTMSERMENT   DRILITERALKINESSOMTIPDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUDOAYSEMKM   GRYWSKERRGGHTUSKEMMILGSIKKMMKKR   NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF   LSVITTV		į.					OHYOSYMOLIOOKSAVEYAQSQMSLVSMCK
	ì	Į.					
GRYWSKERKQHLVKAKEQRRREFMOSR   LCLKEQQAADDREMMLESHKKMMKKK   NKKIPDNWMTIQELLTHGTKSPDGTRVYNSF   LSVTTV   LSVTTV   LSVTTV   CIGKNGGGAFLIPYELTLIFAGYPLFLLECSLQ   QTSIGGLGWKLAPMFKGVGKAADLP   DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY   LGKNNGGAFLIPYELTLIFAGYPLFLLECSLQ QYTSIGGLGWKLAPMFKGVGLAAVLSFW   LNIVYIVISWAYYLYNSFTTLWKQCDNP   WNTDRCFSNYSMVNTTNNTSAVVEFWERN   MHOMTOGLDKPGQIR WPLAITLAIAWILVYF   CIWKGVGWTGKVVYFSATYPYIMLILIFFRGV   TLPGAKEGILFYITPNFRKLSSEWLDAA7Q   TLPGAKEGILFYITPNFRKLSSEWLDAA7Q   TLPGAKEGILFYITPNFRKLSSEWLDAA7Q   TLPGAKEGILFYTPNFRKLSSEWLDAA7Q   TLPGAKEGILFYTPNFRKLSSEWLDAA7Q   TLPGAKEGILFYTPSHRINVYRSIMV   AASGPGLAFLAYPEAVTQLPISPLWAILFFSM   LLMGIDSOPCTVGGFTALVFWFLDYNS   ASGMSLLFLVFECVSISWFYGNFYDNIOG   MYGSRPCIWWKLCWSFTPIIVAGYFTSAVQ   MTPLTMGNVFFKWGGGVVWLFALSSWML   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFSMYLDAATA   TLPGAKEGILFYTPLFFFKLCSFQVFLSFPLATTVIDVGLVIP   LVKSFNYHVYVLLVLSGLLGTPTFLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLLFNIGLPTPLHFKIRL   AWFEKMICYLQLTFNIGLPTPLHFKIRL   AWFEKMICYLQLTFNIGLPTPLAFAGGSSRENL   ALQNVQARIRMLATAFA   LPPLAGQVSGTOEDLMGMYTAELSVYGKL   LPLAGQVSGTOEDLMGMYTAELSVYGKL   RKVAKMGPYSMYCKLLGMWRHICTPRQVAD   KVYRFFSEISLGPACWLWDFLRRSQQAGFL   LPLAGQVSGTDEDEMGMYTAELSVYGKL   RKVAKMGPYSMYCKLLGMWRHICTPRQVAD   KVYRFFSEISLGPACWLWDFLRRSQQAGFL   LPLAGQVSGTDEDEMGMYTAELSVYGKL   RKVAKMGPYSMYCKLLGMWRHICTPRQVAD   KVYRFFSEISLGPACWLWDFLRRAQPABARMA   RRAAGGAPLSAAAASPPPFOTPRCPVPL   LLLLLGARAGGASTLLVCGSTVGGCQCQRRGNISA   AVRYFPAAPAEPVTVYPSMLNVAANHPNAS   TVGULVPPAAPAEPVTVPSMLNVAANHPNAS   TVGULVPPAAPAEPVTVPSMLNVAANHPNAS   TVGULVPPAAPAEPVTVPSMLNVAANHPNAS   TVGULVPPAAPAEPVTVPSMLNVAANHPNAS   TVGULVPPAAPAEPVTVPSMLNVAANHPNAS   TVGULVPPAAPAE	1	ļ	1	ł	1	ł	I
LDCLKEQQAADDRKEMNILEISHKKMKKR   KKKIPDNWATQELTHGTKSPDGTRYYNSF   LSYTTV	ŀ				İ		
NKKIPDNWMTIQELLTHGTKSPDGTRYNSF							
LSYTTY	ļ.						
725   2075   A   5707   3   1770	ļ						
DRDIWKGRFDFLMSCVGYAGLGLOWURFPY LCCKNGGGAFLIPYETLI-IRAGVPLFILECSLG QYTSIGGLGVWKLAPMFKGVGLAAVLSFW LNIYIVIISWAIYYLYNSFITTLEWKQCDNP WNTDRCFSNYSMYNTTIMTSAVVEFWERN MHQMTDGLDKPGGRWPLATILAIAULLYFF CIWKGVGWFGKVYFSATYPYIMLIILFFRGV TLPGAKEGILFYITPFKRLSDSEVUDAATQ IFFSYGLGLGSLIALGSYNSPHNNVYRDSIIVC CINSCTSMFAGFVTESIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWALIFFSM LLMLGIDSOFCTVEGFTIALVDEYPRLLRINRR ELFIAAVCISY11GLSNITQGGIVYKLFDYYS ASGMSLIFLYFTECVSISWYGVNRFYDNIQE MYGSRPCIWWKLCWSFFTPIIVAGYFISAVQ MTPLTMGNYVFPKWGQGGWGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSTSKEAYI  726 2076 A 5711 156 423 PREDPGRTFELRGSAPKKTGANMPVRRGHVA PQNTFLGTIRKFEGQNKKFIIANAKQNCALI YCNDGFCEMGTGSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASSYFFKLCSFQVFLSFPLATIVIDGLVVPI LVSSPNVHYVYLLLVLSGLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVYVLLVLSGLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVYVLLVLSGLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVYVLLVLASLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVYVLLVLASLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVYVLLVLASLLFYPILHFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYVDFLLRSQQAGFL LPLSGGVDSAATACLIYSMCQVCGAVRSON EEVLADVRTIVNQISYTPQDPRDLCGRLTTC YMASKNSSQETCTRABELAQQIGSHLSINID PAVKAVMGPSSLYTGKLSPAHAGSSRENI ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINFICO IKKTULARAVGCTQCPAQVGCDEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAE LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLADGQVSQTDEEDMGMTYAEATAEA LEPLAD	1	1					
LCGKNGGGAFLIPYELTLIFAGVPLILECSLG QYTSIGGI_GOWKLAPMEGOGIAAAVLSFW LNIYYIVIISWAIYYLYNSFITTILPWKQCDNP WNTDRCFSNYSMYNTINMTSAVVEWERN MHQMTDGLDKFQQR WPLAITLAIAWILLYF CIWKGYGWTGKVVYFSATYFYMILIIFRGY TLPGAKEGILEYIPTPFKRLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSPHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVYRRSIAVV AASGFGLAFLAYPEAVTQLFISFLWAILFFSM LLMLGIDSQFCTVEGFTTALVDEYSFLANTRYDNIQS WOGSRPCIWKLCWSFFPIIVAGVFFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IFGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVQCAUNQSEDIV PQXSLCTCTC  127 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFLAATHQSVCCCI PLYSSPNVHYYVLLULSGLLFYIPLHFKIRL AWFEKMTCYLQLLFNICLPDVSEE PGYMAYMFLIKGSLKQRQVGLAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTIC YMASKNSQETCLARELQQIGSHHSINID PAVKAVMGIFSLVTGKSFLAAHGGSSERNIL ALQNVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EEVLADVRITVNQISYTPQDPRDLCGRILTIC YMASKNSQETCICARELQQIGSHHSINID PAVKAVMGIFSLVTGKSFLAAHGGSSERNIL ALQNVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EVVLADVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EVVLADVRGPSPLTYTTSYQVAD KVKFFSKYSMRNHKMTTLTPAYHANYSPE DNRPILRPFLYNTSVPWQPRCCENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCARLSRARAPGPGAAGAGRKRLADPGP PASRILRAPGSRPRLAFCTRAAQPAHARMA PRAAGGAPLSARAAASPPPQTPPCVPLL LLLLGAARAGALELQRAFSPPTPTNFRALDC AAGTVYLAAVNRLYQLSGANISAV AVRPPAAPPAEPVTVFSSNLNVAAMFINAS TVGLVLPPAAAGAGGSRLUAGATYTGYGSSFF	725	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP
LCGKNGGGAFLIPYELTLIFAGVPLILECSLG QYTSIGGI_GOWKLAPMEGOGIAAAVLSFW LNIYYIVIISWAIYYLYNSFITTILPWKQCDNP WNTDRCFSNYSMYNTINMTSAVVEWERN MHQMTDGLDKFQQR WPLAITLAIAWILLYF CIWKGYGWTGKVVYFSATYFYMILIIFRGY TLPGAKEGILEYIPTPFKRLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSPHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVYRRSIAVV AASGFGLAFLAYPEAVTQLFISFLWAILFFSM LLMLGIDSQFCTVEGFTTALVDEYSFLANTRYDNIQS WOGSRPCIWKLCWSFFPIIVAGVFFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IFGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVMQPSEDIV PGYMAYMFLIKGSLKQRQVMVQPSEDIV PGYMAYMFLIKGSLKQRQVQCAUNQSEDIV PQXSLCTCTC  127 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFLAATHQSVCCCI PLYSSPNVHYYVLLULSGLLFYIPLHFKIRL AWFEKMTCYLQLLFNICLPDVSEE PGYMAYMFLIKGSLKQRQVGLAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTIC YMASKNSQETCLARELQQIGSHHSINID PAVKAVMGIFSLVTGKSFLAAHGGSSERNIL ALQNVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EEVLADVRITVNQISYTPQDPRDLCGRILTIC YMASKNSQETCICARELQQIGSHHSINID PAVKAVMGIFSLVTGKSFLAAHGGSSERNIL ALQNVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EVVLADVQARIRMVLAYLFQQLSLWRQHGCPLAVBON EVVLADVRGPSPLTYTTSYQVAD KVKFFSKYSMRNHKMTTLTPAYHANYSPE DNRPILRPFLYNTSVPWQPRCCENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCARLSRARAPGPGAAGAGRKRLADPGP PASRILRAPGSRPRLAFCTRAAQPAHARMA PRAAGGAPLSARAAASPPPQTPPCVPLL LLLLGAARAGALELQRAFSPPTPTNFRALDC AAGTVYLAAVNRLYQLSGANISAV AVRPPAAPPAEPVTVFSSNLNVAAMFINAS TVGLVLPPAAAGAGGSRLUAGATYTGYGSSFF			1		[		DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
QYTSIGGLGVWKLAPHKGVGGLAAVLSFW LNNYTNIUSMATYLYLNSFTTIPWKQCDNP WNTDRCFSNYSMVNTTNMTSAVVEFWERN MHQMTDGLDKPGQIRWPLAITLAIAMILVYF CIWKGVGWFKAYYFSATYPYMLIILFFRGV TLFGAKEGILFYITPYFKKLSDSSVUDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFATPPAYTOLEPSLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCLISY11GLSNITQGGIVYFRLAPPSAV ASGMSLLFLVFFECVSISWFYGVNRFYDNIGE MVGSRPCIWWKLCWSFFTPIIVAGVFFSAVQ MTPLTMGNYVFFWGGGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEDGPAGSSTSKAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIRKFEGQNKKFILANARQNCAII YCNDGFCEMGTGSRPDVMQKPCTCD  4 ST16 3 274 HASSYFFKLCSFQVFLSFPLATIVIDGLVVIP LVSSPNVHYVYLLLVLSGLLFYPLHFFKRL AWFEKMTCYLQLLFNICLPDVSEE TVSSPNVHYYVYLLVLSGLLFYPLHFFKRL AWFEKMTCYLQLLFNICLPDVSEE EVYLADVRTIVNGISYTPODPRDLCGRLTTC YMASKNSSQFTCTRARELAQQIGSHISLNID PAVKAVMGPISLYTGKSLSHLAAHGSSSRENL ALQNVQARRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVGFCIQRFQLPALGSLLAFATAE LEPLADGQVSQTDEEDMGMTYABLSVYGKL RKVAKMGPYSMFCKLLGPLPALGSLLAFATAE LEPLADGQVSQTDEEDMGMTYABLSVYGKL RKVAKMGPYSMFCKLLGPLPALGSLLTTC YMASKNSSQETCTRARELAQQIGSHLSLNTD PAVKAVMGPSSLFTCARAEQAGGSKINLSNID PAVKAVMGPSSLFTCARAEQAGGSKINLSNID PAVKAVMGPSSMFCKLFAHAGGSSRENL ALQNVQARRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTTCYMASKNSSQETCTRARELAQQIGSLLAFATAE LEPLADGQVSQTDEEDMGMTYABLSVYGKL RKVAKMGPYSMFCKLLGWRHICTPRQVAD KVKFFSKYSMNRHKMTTLTPAYHABNYSPE DNRPDLRPFLYNTSWPWQPRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAFGSSPALAGGSRENL ALQNVQARRMVLAYLFAQLSLKARAPGPAAPPTVPRSNLNVAAANFINAS LLLLLLAAAVAGGAPLSLAFAAVG PVPDSPLCHARPQSCPPPPTPTPNFRLDL LLLLLGAARAGGAPLERATLDNYKKL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFFSNLNVAAHRPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1	1	1	1		
Linyyiviiswaiyyinsettifipwigconp	1	}	1		1	}	
WNTDRCFSNYSMVNTTNMTSAVVEFWERN MHQMTDGLDKPGQRW WPLAITLAIAWILVYF CIWKGVGWTGKVVYFSATYPYIMLIILFRGV TLPGAKEGILFYITPNFRKLSDSEWULDAATQ IFFSYGLGGSLIALGSYNSFHNVYYRDSIDVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTOLISPLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFLAAVCISYLIGLSNITQGGIVYFKLFDYYS ASGMSLLFUFFECVSISWFYGVNRFYDNIQE MYGSRPCIWWKLCWSFFFIILVAGVFIFSAVQ MTPLTMGNYVFFKWGOVIVAGVFFSAVQ MTPLTMGNYVFFKWGOVIVAGVFIFSAVQ MTPLTMGNYPFKWGOVIVAGVFIFSAVQ MTPLTMGNYPFKWGOVIVAGVFIFSAVQ MTPLTMGNYFFKWGOVIVAGVFIFSAVQ MTPLTMGFTPELRGSAPRKTGANNPVRRGHVA PQNTFLGTIIRKFBQONKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASSEYFFKLGSPQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYPLHIFKIRL AWFERMTCVLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATAGLIYSMCCQVCEAVRSON EEVLADVRITIVNQISYTTQDFRDLGGRLTTC YMASKNSQGTCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAHGGSSRENL ALQNQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYTDCSSADINFIGG ISKTDLRAFVQFCQRFQLPALQSILLAPATAE LEPLADGVOSTOTEEDMGNTYAELSYVGKL KVARFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRFFLYNTSWPQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGFGAAGAGRKRLADPGPP PASRRIRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFPTQTPPRCPVPLI LLLLLGAARAGALEIGRRFPSFTFTNNFALDG AAGTYYLAAVNRLYQLSGANISLEAEAAVQ PVPDSPLCHAPQLPQASGIPBRRLTONYNKIL QLDPGQGLVVVGGSIYQGFCQLRRGGNISAV AVRPPAAPPAEPVTVFPSMLNVAANHIPNAS TVGLVLPPAAGAGGSSTLJVGATYTGYGSSFF	1		1				
MHQMTDGLDKPGQIRWPLAITLAIAWILVYF   CIWKGVGWTGKVVYPATYPYMILIIFFRGV   TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ   IFFSYGLGLGSLAIGSYNSFHNNYRDSIIVC    CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV     AASGPGLAFLAYPEAVTQLPISPLWAILFFSM     LLMILGIDSQFCTVEGFILAUDEYPRILIRNRR     LEHIAAVCIISYI IGLSNITQGGIVYPKLEDYYS     ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE     MYGSRPCIWWKLCWSFFTPIIVAGVPIFSAVQ     MTPLTMGRYVPFKWGQGVGWMLALSSMYJ     IFGYMAYMFITLKGSLKQRIQVMVQPSEDIV     RPENGPEQPQAGSSTSKEAYI     PRENGPEQPQAGSSTSKEAYI     PRENGPEQPGAGSSTSKEAYI     PRENGPEQPGAGSSTSKEAYI     PRENGPEQPGAGSSTSKEAYI     PRENGPEQPGAGSSTSKEAYI     ASSYFFKLCSFOVPLSFPLATIVIDVGLVVIP     LVKSPNVHYLVYVILLLVLSGLLFYPLIHFKIRL     AWFEKMTCYLQLLFNICLPUSPEL     ASSYFFKLCSFOVPLSFPLATIVIDVGLVVIP     LVKSPNVHYLVYVILLLVLSGLLFYPLIHFKIRL     AWFEKMTCYLQLLFNICLPUSPEL     WKYHSPEEISLOPACWLWDFLRSQQAGFL     LPLAGDGVSATAGLIYMCQVCEAYRSON     EVLADVRTIVNOISYTPQDPRDLGGRITTC     YMASKNSSGTCTRARELAQGIGSHHISLNID     PAYKAVMGIPSLVTGKSPLFAAHGGSSRENI     ALQNVQARIRMVLAYLFAQLSLWSRGVHGG     LLVLGSANVDESLLGYKTYDYCSSSADINPIGG     ISKTDLRAFVQFCIQRFQLPALQSILLAPATAB     LEPLADGVSCTDEEDMMTYABLSVYGKL     RKVAKMGPYSMFCKLLGMWRHICTPRQVAD     KVKRFSKYSMNRHKMTTLTRAYHAENYSP     A 5741 1 5976     PGCAARISARAPQFGAAGAGRKRLADFOPP     PASRRLRAPGSRPRLAPCTRRAQPAHARMA     PRAAGGAPLSARAAAASSPPFTCPPRCCPVLL     LLLLLGARRAGALEIGRRFFSFTFINNFALDG     AAGTYYLAAVNRLYQLSGANISLEAEAAY     PROPRIDE TO THE TOTAL TO THE TOTAL THE TOT		}			i		
CCIWÄGOYGHTÜK-VVPSATYPYIM.IILEFRGY TLPGAKEGILFYITPINFRKLSDSEVWLDAATQ IFFSYGLGLGSLALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIRSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRIRR ELFIAAVCIISYI IGLSNITQGGIYVFKLFDYYS ASGMSLLELVFFECVSISWFYGVYRFYDNIQE MYGSRPCIWWKLCWSFFIPIVAGYFISAVQ MYPLTMGNYVFPKWQQQYGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTTLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMGKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSFNVHYVYLLLVILSGLIFYPLHFKIRL AWFEKMTCYLQLLFNICLIPFDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WXYHSPEEFISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPDLCGRELITC YMASKNSQETCTRARELAQQIGSHHISLNID PAVKAVMGIPSLVTGKSPLFAAHGGSSRENL ALQNOARIRMVLAFPAQLSLWSRGVHGG ISKTDLRAFVQFCIQRFQLFALGSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSYVGKL RKVAKMGPYSMFCKLLGMWHICTPRQVAD KVKFFSYYSMRHKMTTLFPAYHAENYSPE DNRFDLRPFLVNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCARALSRARAPGFGAAGAGRKRLADPGPP PASRRLRAPGSRPLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEJGRRFSPTFTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PYDSPLCHAPQLPQASCEHPRILTDNYNKLL QLDPGGGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAEDLAPACPQASCEHPRILTDNYNKLL QLDPGGGSLLVVCGSIYQGFCQLRRRGNISAV AVRFPPAEPVTVFPSMLNVAANHFNAS			1		1	1	
TLPGAKEGILFYITPNFRKLSDSEWNLDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGFGLAFLAYPEAVTQLIPSILWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS ASGMSLLFUFFECVSISWFYGVNRFYDNIQE MYGSRPCIWKLCWSFFTPIIVAGVFIFSAVQ MYPLTMGNYVFPKWGQGYWLMALSSMVL IPGYMA YMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGONKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLGSFQVFLSFPLATIVIDVGLVVIP LVKSSNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLFDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSON EEVLADVRTIVNQISTYPOOFROLGGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINFIGG ISKTDLRAFVQFCTQRFQLPALOSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTFRQVAD KVKRFFSSYSMRMKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVU  729 2079 A 5741 1 5976 PGCAARLSRARAGFGAAGAGRKRLADPGPP PASRRLRAPGSRPLAPCTRRAQPAHARMA PRAAGGAPLSARAAASPPPFOTPPRCYPLL LLLLLGAARAGALEIQRRFFSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PYDSPLCHAPQLPQASCEHPRILTDNYNKIL QLDPGGGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPAEPVTVFPSMLNVAANHPNAS TVGLUVLPPAAGAGGSRLLVGATYTGYGSSFF	}	1	l	ļ	<u> </u>		
IFFSYGLGLGSLIALGSYNSFIRNNVYRDSIIVC   CINSCTSMFAGFVTFSIVGFMAHVTKRSIADV   AASGPGLAFLAYPEAVTQLPISPLWALLFRSM   LLMLGIDSQFCTVEGFTTALVDEYPRLLAFRING   ELFIAAVCIISY11gLSNITQGGIYVFKLFDYYS   ASGMSLLFVFFECVSISWFYGVNRFYDNIQE   MYGSRPCIWWKLCWSFFFIPIVAGVFIFSAVQ   MYPLTMGNYYFFKWGQGVGWLMALSSMVL   IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV   RPENCPEQPQAGSSTSKEAY   PREPROPEQPQAGSSTSKEAY   PREPROPEQPGAGSSTSKEAY   PREPROPEQPGAGSSTSKEAY   PREPROPEQPGAGSSTSKEAY   PREPROPEQPGAGSSTSKEAY   PREPROPEQPGAGSSTSKEAY   PREPROPEGPGAGSFTSKEAY   PREPROPERTYPLATIVIDVGLVVIP   LVKSPNYHYVYLLLVLSGLFYIPLHFKIRL   AWFEKMTCYLQLLFNICLPDVSEE   LVKSPNYHYVYVLLLVLSGLFYIPLHFKIRL   AWFEKMTCYLQLLFNICLPDVSEE   WKYHSPEEFISLGPACWLWDFLRRSQQAGFL   LPLSGGVDSAATACLTYSMCCQVCAAYSGN   EEVLADVRITVNQISYTPQDPRDLCGRILTTC   YMASKNSSQETCTRABLAQAGSHHISLNID   PAVKAVMGIFSLYTRSVCQVCAAYSGN   EEVLADVRITVNQISYTPQDPRDLCGRILTTC   YMASKNSSQETCTRABLAQAGSSENLL   ALQNVQARIRMVLAYLFAQLSLWSRGVHGG   LLVLGSANVDESLLGYLTKYDCSADINFIGG   ISKTDLRAFVQFCQRPQLPALQSILLAFATAE   LEPLADGQVSQTDEEDMGMTVAELSVYGGL   RKVAKMGPYSTRYKMKMTITJPAYHAENYSPE   DNRFDLRPFLYNTSWPWQFKCIENQVLQLER   AEPQSLDGVD   A 5741   5976   PGCAARLSRARAGGSPGAAGAGGKRLADPGPP   PASRRLRAFGSRPLAPCTRAAQPAHARMA   PRAAGGAPLSARAAASPPFFQTPPCVPLL   LLLLGAARAGALEIQRRFPSPTPTNNFALDG   AAGTVYLAAVNRLYQLSGANLSLEEAAAVG   PVDSSPLCHAPQLPQASCEHPRR.TDNYNKLI   QLDPGGGLVVVCGSIYQGFCQLRRRGNISAV   AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS   TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	[	İ	1			
CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWAILFFSM LLMLGIDSQFCTVGGFITALVDEYPRLLRNRR ELFIAAVCIISYI IGLSNITQGGIVYFKLFDYYS ASGMSLIFLVFECVSISWFYGVNRFYDNIQE MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMQVPSEDIV RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNITLGTIIRKFEGQNKKFILANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYPFKLCSPYLSPFLATVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASKASPYPRVVPFLSPFLATVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDAATACLTYSMCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQGISHHISINID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENI ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYYLTKYDCSSADINPIGG ISKTDLRAFVGPCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHCTTPRQVAD KVKRFFSKYSMRRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWGPRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRRRLAACASPPPFOTPPRCVPVLL LLLLLGARAGGAELSRAAAASSPPFTTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQDQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGRNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAAGAGGSRLLVGATYTGYGSSFF		1			1	1	TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ
AASGPGLAFLAYPEAVTOLPISPLWALIFFSM LLMIGIDSQFCTVEGFITALVDEYPRILRNRR ELFIAAVCIISY1 IGLSNITQGGIYVFKLFDYYS AASGMSLLFLVFFECVSISWYGVNRFYDNIQE MYGSRPCIWWKLCWSFTPIUVAGVFIFSAVQ MYBLTIMGNYVFPKWGQGVGWLMALSSMYL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSTSKEAY1  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYYVLLLVLSGLLFYPLIHFKIRL AWFEKMTCYLQLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSON EEVLADVRTIVNQISYTPQDPRDLCGRLTTC YMASKNSSQETCTRARELAQQIGSHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMRHKMTTILTPAYHAENYSPE DNRFDLRPFLYNTSWPWGPRCLENQVLQLER AEPGSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRRRLAAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFTQTPPRCVPULL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRILYGLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAAGAGSSRLLVGATYTGYGSSFF		1				1	IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
AASGPGLAFLAYPEAVTOLPISPLWALIFFSM LLMIGIDSQFCTVEGFITALVDEYPRILRNRR ELFIAAVCIISY1 IGLSNITQGGIYVFKLFDYYS AASGMSLLFLVFFECVSISWYGVNRFYDNIQE MYGSRPCIWWKLCWSFTPIUVAGVFIFSAVQ MYBLTIMGNYVFPKWGQGVGWLMALSSMYL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSTSKEAY1  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYYVLLLVLSGLLFYPLIHFKIRL AWFEKMTCYLQLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSON EEVLADVRTIVNQISYTPQDPRDLCGRLTTC YMASKNSSQETCTRARELAQQIGSHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMRHKMTTILTPAYHAENYSPE DNRFDLRPFLYNTSWPWGPRCLENQVLQLER AEPGSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRRRLAAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFTQTPPRCVPULL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRILYGLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAAGAGSSRLLVGATYTGYGSSFF	ł		ł	ł	1		CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCIISYI.IGLSNITQGGIYVFKLFDYYS ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE MYGSRPCIWKLCWSFFTPIIVAGYFIFSAVQ MYDLTMGNYVFPKWGQGVGWLMALSSMVL PGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNITLGTIRKFEGQNKKFILANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCQOVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIPSLVTGKSPLFAHAGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHCTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPTQTTPRCPVPLL LLLLLGAARGALEIQRRFSPIPTINNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHERRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAARGAGGSRLLVGATYTGYGSSFF		1	ļ	1	1	}	I I
ELFIAAVCIISYI-IGLSNITQGGIVVFKLFDYYS	į	1	l	İ		1	
ASGMSLJELVFFECVSISWFYGVNRFYDNIQE   MVGSRPCIWKLCWSFFTPIIVAGVFIFSAVQ   MTPLTMGNYVFFKWGQGVGWLMALSSMVL   IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV   RPENGPEQPQAGSSTSKEAYI   PRRDPGRTPELRGSAPRKTGANMFVRRGHVA   PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII   YCNDGFGEMTGSRPDVMQKPCTCD   YCNDGFGEMTGSRPDVMQKPCTCD   YCNDGFGEMTGSRPDVMQKPCTCD   LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL   AWFEKMTCYLQLLFNICLPPVSEE   AWFEKMTCYLQLLFNICLPPVSEE   UKYHSPEEISLGPACWLWDFLRRSQQAGFL   LPLSGGVDSAATACLIYSMCCQVCEAVRSGN   EEVLADVRITVOJSYTPOJPRDLCGRILTTC   YMASKNSSQETCTRARELAQQIGSHHISLNID   PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL   ALQNVQARIRVLAYLFAQLSLWSRGVHGG   LVLGSANVDESLLGYLTKYDCSSADINPIGG   ISKTDLRAFVQFCQRRQLPALQSILLAPATAE   LPPLADGGVSQTDEEDMGMTYAELSVYGKL   RKVAKMGPYSMFCKLLGMWRHICTPRQVAD   KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE   DNRFDLRPLVNTISWPWQFRCIENQVLQLER   AEPQSLDGVD   PGCAARLSRARAPGFGAAGAGRKRLADPGFP   PASRRLRAPFSRPLAPCTRAAQPAHARMA   PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL   LLLLLGAARAGALGRAFSSPTIPTNNFALDG   AAGTVYLAAVNRLVQLSGANLSLEAEAAVG   PVPDSPLCHAPQLPQASCEHPRRLITDNYNKIL   QLDFGQGLVVVCGSIYQGFCQLRRRGNISAV   AVRPPPAAPPAAPPAEPVTVFPSMLNVAANHPNAS   TVGLVLVPPAAGAGGSRLLVGATYTGVSSFF	1.			1		1	
MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ MTPLTMGNYVPFKWQQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI  726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIRKFEGONKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSPQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMITCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAYKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG ILVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRTDLRPPLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRILRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPITPINFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDFQQGLVVVCQSIYQGFCQLRRRGNISAV AVRFPFAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	1	1	1		1	
MTPLTMGNYVFPKWGQGVGWLMALSSMVL   IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV   RPPNGPEQPQAGSSTSKEAYI     726				1		1	
PGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI		l .	1	1			
726 2076 A 5711 156 423 PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGGAPKTGANMPVRRGHVA PQNTFLGTIIRKFEGGAPKTGANMPVRRGHVA PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLFYIPLIHFKIRL AWFEKMTCVLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGQVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTIDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASSPPPTQTPPRCVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRLTDNYNKIL QLDPGQGLVVVCGSIYQGFQLRRRGNISAV AVRFPPAAPPAEPVTYPSSMLNVAANHFINAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	ł	1		1			
T26					1		IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV
PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLHIFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENI. ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLFALGSILLAFATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRFFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHFNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	ļ	l .	1	1		ł	RPENGPEQPQAGSSTSKEAYI
PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD  727 2077 A 5716 3 274 HASEYFFKLCSPQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSYVGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFFLATIVIDVGLVVIP LVKSPNVHYVYVLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCVPLL LLLLLGAARAGALEIQRRFPSPTPTINNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLITDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRPPAAPAAPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF					"""	)	
727 2077 A 5716 3 274 HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYYYVLLLVLSGLLFYPLHFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNISSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRFPSFTPTINNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRPPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	l.	ĺ				
LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRWVLAYLFAQLSLWSRGVHGG ILVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASSPPPGTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	707	0077	-	6716	<del> </del>	274	
AWFEKMTCYLQLLFNICLPDVSEE  728 2078 A 5737 1899 649 IQASRASPYRRVVDFALSCHEDLLAPISEPIE WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG ILVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPLAPCTRRAAQPAHARMA PRAAGGÁPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	727	2077	A	3/16	3	2/4	
728 2078 A 5737 1899 649 IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNISSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGGSRLLVGATYTGYGSSFF			į .	i		ł	
WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGFGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGÅPLSARAAAASPPFQTPPRCPVPLL LLILLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGGSRLLVGATYTGYGSSFF			<u>L</u> .			<u> </u>	
WKYHSPEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRITVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGFGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGÅPLSARAAAASPPFQTPPRCPVPLL LLILLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGGSRLLVGATYTGYGSSFF	728	2078	Α	5737	1899	649	
LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	1	l	1	]	1	
EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	]	1		Į.	I	
YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFQTPPRCPVPLL LLLLLGAARAGALEIQRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1	1	1	1			
PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1	1	1			
ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1	l	1	1		
LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPFFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1 .		1		
ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1		1			
LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1		1		1.	
RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	-	]		1		1	
RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	i	1					LEPLADGQVSQTDEEDMGMTYAELSVYGKL
KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1	1			
DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	1		1				
AEPQSLDGVD  729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1	1	1			
729 2079 A 5741 1 5976 PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF						1	,
PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1-22	<del> </del>	<del> </del>	<del> </del>	L	
PRAAGGAPLSARAAAASPPFFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	729	2079	[ A.	5741	1	3976	
LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	i	1	1	1			
LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1				
AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF		1	1			1	
PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF	ľ	1	1	1	1		
QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1	[.	1 .		
AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1	1	1		
TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF			1		İ		
• • • • • • • • • • • • • • • • • • •	1	1					1
PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT	1			1		1	
	L	L_	<u> </u>	<u> </u>	L		PRNKSLEDHRFENTPEIAIRSLDTRGDLAKLFT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histdine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, v=possible nucleotide insertion  FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHP SDPPPGAQSYAYLALNSEARAGDKESQARSL LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRLVSVFPARERLFAVFERPQGSPA ARAAPAALCAFRFADVRAAIRAARTACFVEP APDVVAVLDSVVQGTGPACERKLNIQLQPEQ LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV AVASVNNYTAVFLGTVVNGRLLKINLNESMQ VVSRRVYTVAYGEPVHHVMQFDPADSGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ LAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEIFEA VWVNESVVRCDQVVLH TTRKSQVFPLSLQLKGRARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAPERAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLITUHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVQ LPITIQVGNFNQTLATLQLGGSETAIIVSIVICSV LLLLSVAALFVFTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGGWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFVHALEQQKDFAVRD RCSLASSLLDNKHPLIVFYHDELAPKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSREDKPPLAVKYFFDFLEEQAEKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDDI
730	2080	A	5744	3	292	QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD
731	2081	A	5747	1	382	FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
eotide seq-	seq- uence	ŀ	09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciicc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
40,100		ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
		ł	ł	<u> </u>	1	LSFSLRSSRVSGRHWKNFALVPLLREASARD
						RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS GEK
732	2082	A	5753	198	3	AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP PGKLQAQLPCPSPVRFTSARIPPASRPQTKS
733	2083	Α	5754	2	2223	AAGPPGLEAEGRAPESAGPGPGGDAAETPGL
		1				PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL
						SVANCLGAQTVQAPAEPAAGKAEQGETSGR
į				Ì		EAPEAPAVGREDASAEDSCAEAGASGAADG
				1	[	ATAPKTEEEEEEEETAEVGRGAEAEAGDLEQ LNRTSTSTKSAKSGSEASASASKDALQAMILS
						LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN
			İ	1	1	LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK
				ļ		GLSROMIGEFLGNSKKQFNRDVLDCVVDEM
				ŀ		DFSSMELDEALRKFQAHIRVQGEAQKVERLIE
						AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL
1						NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG
]	1	}		-		ADIPRELVVGIYERIQQKELKSNEDHVTYVTK
					}	VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV NKLQKQAAHQREVFLFNDLLVILKLCPKKKS
						SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV
		1				TPLSGSEKKOVLHFCALGSDEMOKFVEDLKE
]		ļ	1			SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA
						QGDPQSKQGSPTAKRFAALRERPAESTVEVSI
		ļ	1	ļ		HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR
			1			QQTPPLPPPPTTPPGTLVQCQQIVKVIVLDKPC
}					1	LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG
						SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH GHRYSSGSRSLV
734	2084	A	5788	8	362	SSVMGDLVGQGLEEQIVARDENSWLIDGGTP
134	2004	^	3786	١	302	IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR
				1		KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT
						RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL
				1		PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST
		1		1		AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG
			1	l		NLVVCLMVYQKAAMRSAINILLASLAFADM LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF
1		Į		1		FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR
		1				AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA
1	ł	1		I		PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY
				1		SFMGILNTLRHNALRIHSYPEGICLSQASKLGL
		1	1			MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC
						WAPFTTYSLVATFSKHFYYQHNFFEISTWLL
1	1	1	ł	l		WLCYLKSALNPLIYYWRIKKFHDACLDMMP
		<u>                                     </u>	1	<u> </u>	1000	KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP
		Į.				LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR
'''	2007	1	30/1		321	RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH
				1		GLARRLTTARRPPASSEQAQQELFNELKPAV
			1			DGANFIVNHMRDQNNYNEEKDSWNRVART
l	1	ł	1		1	VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP
L		<u> </u>				QPFPGDPYSYNVQDKRF1
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG
		1				LLDSSKLCDYENRFNTSKGGELPDRPAGVGV
l	}	l			}	YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS
L	<u>L</u>			J	<u>L</u>	MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  WCSQGADCITPGLYAMVGAAACLGGVTRMT VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL AMDVMKPRRNDPLLTVLTQDSMTVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI LFN
739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP DQALQELRKVARINGHKEAKNLTIEVLMSSV KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA VDFLGRATTALLLSFLGRRTIQAGSQAMAGL AILANMLVPQDLQTLRVVFAVLGKGCFGISL TCLTIYKAELFPTPVRMTADGILHTVGRLGA MMGPLILMSRQALPLLPPLLYGVISIASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM
741	2091	A	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV
742	2092	Α	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC
743	2093	A	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRIL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
744	2094	A	5966	149	327	LVMEELKETTNSSEI SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
745	2095	A	5970	413	856	LYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL NFPRRAQGPTWRSLEVIQSTEL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence	-	/=possible nucleotide deletion, \=possible nucleotide insertion
746	2096	A	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS RCARHGACQRSCLASQDPYCGWHSSRGCVDI RGSGGTDVDQAGNQESMEHGDCQDGATGSQ SGPGDSAYGVRRDLPPASASRSVPIPLLLASV AAAFALGASVSGLLVSCACRRAHRRGKDIE TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP APALLGGPSPRPHECASPLRLDVPPEGRCASA PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG GRFNF
747	2097	A	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI LKLEQENCTLVTTFRGHTGGVTALCWDPVQ RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVVWNMD VERQETPEWLDSDSCQKCDQPFFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDEERAPTATFHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS
748	2098	A	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRRFESPLLWQSAIMIL TMLLMLKLCTEVRVANELNARRRSFTAADS KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
749	2099	A	6002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD
750	2100	A	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC
751	2101	A	6007		1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK
752	2102	A	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of NO: of NO: of nucl-peptide cotide sequence   USSN period of peptide sequence   Sequ	LLHLNG STIVTKW GIIGGHV APTFPV RPNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN MQPISQ SKDMTP LGHCDF
nucl- eotide seq- uence    Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide seq- uence   Deptide sequence   Deptid	LLHLNG STIVTKW GIIGGHV APTFPV APTFPV APTSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
eotide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 corresponding to last amino acid residue of peptide sequence  USSN 09/496 PUS	LLHLNG STIVTKW GIIGGHV APTFPV APTFPV APTSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
sequence    Sequence	LLHLNG STIVTKW GIIGGHV APTFPV APNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
uence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  915  107  107  107  107  107  107  107  1	LLHLNG STIVTKW GIIGGHV APTFPV APNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
amino acid residue of peptide sequence	LLHLNG STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor /=possible nucleotide deletion, \=possible nucleotide insertion  LLRISIPLDSNLRPEKCRFVHPQWQ TIHSTSEADTEPCVDGWVYDQSYFPSDLVCDYQSLKSVVQFLLLTGMLVGG SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITIN KALRKVARTNGIKNAEETLNIEVVRSLDAAQTKTTVWDLFRNPSMRKRICH KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLN SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	LLHLNG STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion  LLRISIPLDSNLRPEKCRRFVHPQWQITHSTSEADTEPCVDGWVYDQSYFPS DLVCDYQSLKSVVQFLLLTGMLVGG SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIT VVLSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNIKALRKVARTNGIKNAEETLNIEVVR. LDAAQTKTTVWDLFRNPSMRKRICIKNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYPFKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLV. SQAEIRKQILGSSSSGKFFCLYTEEFALKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNIFGIIMSEFPLDPQLSKSILASCEFDCV. AMVTAPNCFSHVPHGAEEAALTCW. GDHFTLISIYKAYQDTTLNSSSEYCV.	LLHLNG STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
sequence    Sequence   Nucleotide insertion	LLHLNG STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
LLRISIPLDSNLRPEKCRRFVHPQWQ TIHSTSEADTEPCVDGWVYDQSYFPS DLVCDYQSLKSVVQFLLLTGMLVGG SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVRS LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSGSVRFFUDVGVERRKVYNPRIRANSLS SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCWGDHFTLISIYKAYQDTTLNSSSEYCV	STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
TIHSTSEADTEPCVDGWVYDQSYFPS DLVCDYQSLKSVVQFLLLTGMLVGG SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVRS LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLY SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	STIVTKW GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
DLVCDYQSLKSVVQFLLLTGMLVGC SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVR: LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLY SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	GIIGGHV APTFPV PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
SDRFGRRFILRWGLLQLAITDTCAAF YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVR: LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLY SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	APTFPV RPNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
YCVLRFLAGFSSMIIISNNSLPITEWIF VVILSSGALNIGQIILGGLAYVFRDW ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVRS LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLN SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	PNSKAL QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
VVILSSGALNIGQIILGGLAYVFRDW. ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVR. LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG. SVRFVIDVGVERRKVYNPRIRANSLN SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW. GDHFTLISIYKAYQDTTLNSSSEYCV	QTLHVV KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
ASVPFFVFFLLSRWLVESARWLIITNI KALRKVARTNGIKNAEETLNIEVVR: LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLN SQAEIRKQILGSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	KLDEGL STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN /MQPISQ SKDMTP LGHCDF
KALRKVARTNGIKNAEETLNIEVVR: LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	STMQEE LVFLRK CEQDIEK KEKCSL EFLIWSN MQPISQ SKDMTP LGHCDF
LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	LVFLRK  CEQDIEK KEKCSL EFLIWSN MQPISQ SKDMTP LGHCDF
LDAAQTKTTVWDLFRNPSMRKRICI KNLKEKA  753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG: SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW: GDHFTLISIYKAYQDTTLNSSSEYCV	LVFLRK  CEQDIEK KEKCSL EFLIWSN MQPISQ SKDMTP LGHCDF
753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	CEQDIEK KEKCSL EFLIWSN MQPISQ SKDMTP LGHCDF
753 2103 A 6043 1 1470 DSFESILRLIFEIHHSGEKGDIVVFLAG VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	KEKCSL EFLIWSN /MQPISQ .SKDMTP LGHCDF
VCETVYQGSNLNPDLGELVVVPLYP FKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	KEKCSL EFLIWSN /MQPISQ .SKDMTP LGHCDF
FKPLDETEKRCQVYQRRVVLTTSSG SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	EFLIWSN /MQPISQ .SKDMTP LGHCDF
SVRFVIDVGVERRKVYNPRIRANSLV SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	/MQPISQ SKDMTP LGHCDF
SQAEIRKQILGSSSSGKFFCLYTEEFA LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	SKDMTP LGHCDF
LKPAEMQEANLTSMVLFMKRIDIAG MNRPAPESLMQALEDLDYLAALDN FGIIMSEFPLDPQLSKSILASCEFDCV AMVTAPNCFSHVPHGAEEAALTCW GDHFTLISIYKAYQDTTLNSSSEYCV	LGHCDF
MNRPAPESLMQALEDLDYLAALDNI FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCWI GDHFTLISIYKAYQDTTLNSSSEYCV	
FGIIMSEFPLDPQLSKSILASCEFDCVI AMVTAPNCFSHVPHGAEEAALTCWI GDHFTLISIYKAYQDTTLNSSSEYCV	
AMVTAPNCFSHVPHGAEEAALTCW. GDHFTLISIYKAYQDTTLNSSSEYCV.	
GDHFTLISIYKAYQDTTLNSSSEYCV	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
FGSKENTLNIKKALLSGYFMQIARDY	
YLMLTHKQVAQLHPLSGYSITKKMF	
HKFSISENNYIRITSEISPELFMQLVPC	
PPSESKDILQQVVDHLSPVSTMNKEC	QMCEI
CPETEQRCTLQ	OVOCOUP.
754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMNN	
LMLINVDRYAAIVHPLRLRHLRRPR	
LGVWALILVFAVPAARVHRPSRCRY	
LCFESFSDELWKGRLLPLVLLAEALO	FLLPLA
AVVYSS	
755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQL	
RNARSVKITKRFTKLLIAPESAAPEE	
PEPGRARRSDTHTFNRLFRRDEEGRI	
QGPAGIGKTMAAKKILYDWAAGKL	
FAFFMPCGELLERPGTRSLADLILDQ	
PVPQMLAQPQRLLFILDGADELPALO	
· CTDPFEAASGARVLGGLLSKALLPT	
RAAAPGRLQGRLCSPQCAEVRGFSD	KDKKK
YFYKFFRDERRAERAYRFVKENETL	
PFVCWIVCTVLRQQLELGRDLSRTSI	
LLFITSVLSSAPVADGPRLQGDLRNL	
GVLGRRAQFAEKELEQLELRGSKVQ	
KELPGVLETEVTYQFIDQSFQEFLAA	
DGGVPRTAAGGVGTLLRGDAQPHS	
RFLFGLLSAERMRDIERHFGCMVSE	
LRWVQGQGQGCPGVAPEVTEGAKG	
PEEEEEGEEPNYPLELLYCLYETQED	
LCRFPELALQRVRFCRMDVAVLSYC	
GQALRLISCRLVAAQEKKKKSLGKR	
GG GG	~4. 1010
756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCC	CODAMO
756 2106 A 6060 12 436 SGRPTRPARPTQQMGRPMLTEVCC	
LSHIPGQAFSGLYSLKILMLHNNQLO	
LWELPSLQSLRLDANLISLVPERSFE	JUSSLKH
LWLDDNALTEIPS  LTDLGI GAADNGA FINAL LAND OF THE COLUMN AND A COLUMN AN	CCODET
757 2107 A 6063 54 419 ITPLGLGAADMCAFFWLLLLLLQE	
WRWCGSEEVVAVLQESISLPLEIPPE	
WSSHKSLATVVPGKEGHPATIMVTY	IFH Y QG

			1			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
		1		residue of	sequence	/=possible nucleotide deletion, \=possible
				peptide	1	nucleotide insertion
	<del> </del>		<del> </del>	sequence		QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	A	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
/38	2100	^	0000	123	430	PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
		ì	1	i		LRAVLKLMSECWAHNPASRLTALRIKKTLAK
		l				MVESQDVKI
759	2109	A	6072	3	650	PGRRFRPAALEERAMEKLREKVPFQNRGKGT
139	2109	^	0072	3	030	LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
		ŀ	1			SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA
ļ i			1			LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP
			İ			KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV
i			1			HSSVADMONMPAAVHALLTOPSLSAAPFAQ
		l	l			RYLGTLPSTGSTTLPQCHAGNATVW
760	2110	A	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG
′°°	12110	''	33//	-	1 '33	LRGCMLIELPLRGRLQLEACGMRRKSLLTRK
ļ		ļ	1		ļ	VICKSDAPTGDVLLDEALKHVKETQPPETVQ
		İ				NWIELLSGETWNPLKLHYQLRNVRERLAKNL
						VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR
				]		LIKKVQEAVLDKWVNDPHRMDRRLLALIYL
		1		İ		AHASDVLENAFAPLLDEQYDLATKRVRQLLD
		1	1	l		LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
'''		1	****	555	373	HOKLSADMADHSNLIRSLLVGAEDARLMRD
ļ	1		1			MKTMKSRYMELYDLNRDLLNGYKIRWNNH
						TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
						ACRDAIRSNNINTLFKIMRVGTASS
762	2112	A	6079	2	2686	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS
		į .				HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG
		İ	Ì			SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI
		1	1			QLAKLDSEAWPGVLDSERDRLILINEKEELLK
		1				EMRFISPRKWTQGEVEQLEMARKRLEKDLQ
<b>!</b>		]	]	J	]	AARDTQSKALTERLKLNSKRNQLVRELEEAT
	1					RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR
			<b>!</b>			GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ
	ł					SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ
	Ì	i	1			KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS
		ŀ			i	PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA
		į	1			TLCELSLGNSAQERYRLEEPGTEGKQLGQAV
		l	1			NTAQGCGLKVACVSAAVSDESVAGDSGVYE
<b>J</b>			J	J	ļ	ASVQRLGASEAAAFDSDESEAVGATRIQIALK
						YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR
ŧ			1	[	ŀ	VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW
]				1	!	VSMSYPALHQKTLRVDVCTTDRSHLEECLGG
}	1	l	1	1	1	AQISLAEVCRSGERSTRWYNLLSYKYLKKQS
			1			RELKPVGVMAPASGPASTDAVSALLEQTAVE
	1			1		LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
		1	1			EEEEEVEEEGEEDVFTEKASPDMDGYPALK
		1	1	1		VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF
[	[	1	1	1		LRGSTHRSKTFSPGPQSQYVCRLNRSDSDSST
}		i				LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK
	1	1		1	1	SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS
				1	1	
					`	VLKELKEQLEQAKSHGEKELPQWLREDERFR
					,	LLLRMLEKRMDRAEHMGELQTDKMMRAAA
						LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1					LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV
763	2113	A	6082	3	1558	LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
763	2113	A	6082	3		LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV PHPIRFSKLCVSFNNQEYNOFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED
763	2113	A	6082	3		LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG
763	2113	A	6082	3		LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV PHPIRFSKLCVSFNNQEYNOFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED

السال السال النسب حبيدة في سيند منت

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				·		THVTLHGTELCDESYPALLTDIPVGDLHPGEQ LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF EHLERVYADIPFLLMTDLLSASPWALTIVSSE LHLAPSMTTVDQLESQVDNVILQTGESASECF CLQCPSLGNIEGGVATGHYIISWKRTSAMENI PIITTVITLPHVIVENIPLHVNADLPSFGRVRES LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM DDTSIAAA
764	2114	A	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP PPPAPAPPTLGNNHQESPGWRCCRPTLRERN ALMFNNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA AFLILLKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA
765	2115	A	6099			SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK
766	2116	A	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	A	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY KHEDLQTDESSMDDRHPRRQLCGGNQAATE

			T- 252			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in .	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	}	l	ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
	j	l	j	sequence		nucleotide insertion
						RIILFGRELQALSEQLGREYGKNLAHTEMLQD
ł	l	l	ł	}		AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
		1	1			NSAILESQNLPKQPPLMLALGQASECLRLMA
		ĺ	Ì			RAGLGSCSFARVDDYLH
770	2120	Ā	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
} ```				i -	1	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
	l	1	ļ			VPQEMQMLREVLATLGLGASALANTLAFAH
Ì		1	l .			GNEVIPTIIRARAMGINATFANIAGALAPLMM
	1		1			ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
		ľ	l	1 .	Ì	PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	1	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
1 / / 1	2121	A	0120	909	323	RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
i	Ì	l	1			LTKEDTGWYWCGIQRDFARDDMDFTELIVT
l				<u> </u>		DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
]		}	}	}	1	RKADRSRTSILIICILITGLGIISVISHLTKRRRS
		1	<u> </u>			QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	Α	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
		ļ				TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
	}	1	1	1		WYLLRKHWIANNLFGLAFSLNGVELLHLNN
				1		VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
		l	1	l		FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
			İ		l	VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
1	l	1	ł	l .	l	GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
		1	1	İ		ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
1	1	1	1			EASASKGLEKKEK
773	2123	A	6161	3	1088	COPMLVTRKNHPKLLLRRTESVAEKMLTNW
1113	2123	Ι ^ .	0101	,	1000	FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
		ļ	ŀ			PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
ł	1	1	ı			NPENENAPEVPVKGLDCDTGTQAKEKLLDA
		1	l l		İ	
ł	1	ł	1	l	l .	AYKGVPYSQRPKAADMDLEWRQGRMARIIL
	l .	1	1			QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
l		<u> </u>				ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
		l			]	SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
ł	ł	ł	l	l	1	LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
ĺ	1	1		1		VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ
ļ	1					ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
1	l			J		NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
	1	1	1	ļ		SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
			]			PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
1		1		1	1	GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
[		1	1	1	1	DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
1		l	1			NHTMVYDGFGPADLRQACAELSLWDHGALA
1	1	l				NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
1	ł	ł			ļ	QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	1 2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
1 ''3	2123	^	1 0171	1	3,2	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
1		1	1	Į.		DMK/KSPEIISRRMTFAL*CYSLTFVRFAHYVO
1				I		1 = 1
[	1	1	i			\PWNWLMLGCHTAVDFDQLISSMPCISHGMT
	\	<del></del>	L	<del> </del>	L	ASASAL PROVINCIA PURA PURA POR CONTROL GUIDA
776	2126	Α	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
1					1	RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
1						QAALRGRIQLREQLPRYLMGETQLCTTRCVL
		1				KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
}		1		1	ŀ	VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
						VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
	1	1				VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
1					1	RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
1		1		1		KKRKRKRIRNRSNPKVLSEKQNAEGE
				4	<u> </u>	1

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1			[ ]	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i	İ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			'	peptide		/=possible nucleotide deletion, \=possible
1			ļ	sequence		nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
1		l	i			FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF
Į						YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ
1						RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
1		ł	ł			AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP
1						PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
						SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
1						LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
		•				NSFRYRR
779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
		ĺ				YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
		l				FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
1		l				MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
				<u></u>	<u> </u>	QSQPMY
780	2130	A	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
		l				TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
						HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
						QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
l		l				AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
			ĺ			S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
1			1			FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
						LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
1		ļ				PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
1	Į	j	]			DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
						PV
781	2131	Α	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
		l.				LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
}			1			VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
						QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
		1				LEFLGDLKGCSELKNFQELITQSALVHPKADV
<u> </u>						WWYCGRPLLGTLPSN
782	2132	Α	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
1		j				EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
		1				DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
						KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
		i				DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
1		l	]			ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
1		l				IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA
						EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
1	·	ŀ				SGSRHRRMEAVRLRKENQIYSADEKRALASF
702	2122	<u> </u>	6205	201	1020	NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRREAAEGLHFLGPPGRVRGQ
1						LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
		1				AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
1		1				AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
		Ì				SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
]		1			!	TTVVFWPAKLQASSRVVMFRFEFWDCGESA
1		[				LKKFDHMLLACMENTDAFLFLPSFTDRASFE
						DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
704	2124		(200	06		DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP
1		İ				GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
1		1				PRKPEGAQARTAQSGALRDVSEELSRQLEDIL
						STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
			<b> </b>			TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR
		l			· ·	QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
		<b>J</b>	J .			QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ
L	L	L	L	L	L	KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RSKLESLCRELQRHNRSLKEEGVQRAREEEE KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR QENMELAERLKKLIEQYELREEHIDKVFKHK DLQQQLVDAKLQQAQEMLKEAEERHQREKD
						FLLKEAVESQRMCELMKQQETHLKQQLALY TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI KKLEKETTMYRSRWESSNKALLEMAEEKTV RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR GQRWGSHRTSAVRIFS
785	2135	A	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT KLPWSWGMRPMKIFFSEEYRSISTRISHDAL* EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVRKNENCYNTP GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	6622	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMT\V KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFFGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351		0022	PRSLCFSLWAEAAVLADGGLRRRRILLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

						<del></del>
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		į		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						RQNSEREAGKKHKVEDGTSSVTVLSHEEDA
						MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
			]		J	VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
					[	ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
						I
			t l			EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
			Į .			GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
						KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
		İ			1	CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
		ŀ				DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
		j	j		j	LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
		İ			i	VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
			ļ		ļ	EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
					i	NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
		İ				MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
			ļ		ŀ	KRFEALRATSAEHVNGGLKESEKLSDDLILLL
						QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
					į	EERDCCNDCYLALEHGROFMDNMSGGKVDE
					ļ	ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
				-		LQPVLQDAIQKKRTVRPWGVQGPLTWQOFH
						KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
	ŀ	!				LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
					*	NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
						LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
			1			GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
	1	j	]			SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
					1	ATLASAASSTMTVTSGVAISTSVATANSTLTT
						ASTSSSSSSNLNSGVSSNKLPSFPPFGSMNSNA
						AGSMSTQANTVQSGQLGGQQTSALQTAGISG
	1		1			ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH
			1	ļ		AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
	Į.		ļ		1	GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
	ļ					PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS
				ł	Ì	TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
	1					
'						PPFILAPVKDKQTELGETFGEAGQKYNVLFV
		ļ				GYCLSHDQRWILASCTDLYGELLETCIINIDVP
	1	1			1	NRARRKKSSARKFGLQKLWEWCLGLVQMSS
1	1	1	1		]	LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
	1	1			1	SLSKRLKDMCRMCGISAADSPSILSACLVAM
					1	EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
1	1	J	}		j	NTPQDTSCTHILVFPTSASVQVASATYTTENL
	l '	l				DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
	1	l				NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
	1					RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
1		l			1	LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
l		1			1	QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
	1	l	1		1	ALSWLTCDPATODRRSCLPIHFVVLNOLYNFI
ł	l	l	1	}	1	MNML
789	2120	<del> </del>	6250	1	2002	
107	2139	A	6359	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
				i		LPVGPLLRALATCHALSRLQDTPVGDPMDLK
1	1			1	1	MVESTGWVLEEEPAADSAFGTQVLAVMRPP
1	1				1	LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
	1	1		1	I	SVVVAWPGATQPEAYVKGSPELVAGLCNPET
	1	1				VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
ļ	]	1	Į.	1	]	SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
	1					QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
	1		1			RGCGMVAPQEHLIIVHATHPERGOPASLEFLP
						MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
1				i	1	
		ļ		1		I COTECHNELEDEL I DESIL MOCTUMANICAN
				_	-	LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
				-		LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP EQKTELVCELQKLQYCVGMCGDGANDCGAL KAADVGISLSQAEASVVSPFTSSMASIECVPM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG
						VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR NITDTGFKLLLVGLVTLNFVGGLHAGERARP VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW PPLPAGPLR
790	2140	A	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/L/EKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALGYFHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN
791	2141		6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS FNN\GQLAPGIT\MTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK
792	2142	A	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL
793	2143	A	6446	3201		PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

SEQ ID SEQ ID NO: of	Met S	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	hod I	D NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	I .	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence		09/496 914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence	>	714	ng to first amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			peptide	boque	/=possible nucleotide deletion, \=possible
	ĺ	ļ	sequence		nucleotide insertion
					IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
1 1	. 1				LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
	·				HEEYTRQTAALQQVLSTRILAMKASLCKLSP
					CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
	1				PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN
		ĺ			QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLC\G
1	- 1	1			SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
	}	ŀ			KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
		j			RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
		1			PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
					LQLDGLARAPGQPKIDHLRRLHLGACPTEEC
	- 1	ĺ			KACTRCGCVTMLKSPNRTTAVKQWEQRWIK NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA
		ŀ			APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
	ļ	ŀ			RGPDACVTSRASEEAPAFVOLGPOSTHHSPRT
					PRSLDHLHPEDRP
794 2144	A 6	5490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
705 0145					GSIEPRDLRLQ*AVITPL\TPAWVTQ
795 2145	A 6	5499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
					SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
	ł		ŀ		DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN
		i			STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
	ļ	ł	j		HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
					DTRVRAAVGAPGLPVEPLV
796 2146	A 6	5503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
		1	ĺ		DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
					PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM KKMNKLQNISEELQRNISLQLMSNMNISNKIR
		ł			NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR
		l			WIWHKDSCYFLSDDVQTWQESKMACAAQN
					ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
		Ī			DS/YSWYESG*YNQ\PSAWVIRNAPDLNNMY
1 1 1		1	Ì		CGYINRLYVQYYHCTYKQRMICEKMANPVQ
797 2147	A 6	5507		881	LGSTYFREA PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
214/	^   0	,,,,	•	201	APRAASAQLEERMRDPHPGMTLQEGDCRGS
					QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
		ļ			ESALAKLLLTCCSALRPRATQARGSSRLLVAS
		İ			WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
	.	i	ĺ		AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
.		1			TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
1	ł	ł	ł		ACRISSSDWNTPAPTQSPEEVRRLHLCTSFM
					DMLKALFRTLQAMLLGVWILLLLASLTPLWL /SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
]   . ]	j		J		STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
		[			RAASAQLEERMRDPHPGMTLQEGDCRGSQT
	[	l			VSLTMGTADSDEMAPEAPQHTHIDVHIHQES
	1	Ì		:	ALAKLLLTCCSALRPRATOARGSSRLLVASW
	ł	}	ł		VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
		j			WTGAVAVLAGAAAFIYEKRGGTYWALLRTL
	- 1	ł			LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
	ĺ	ſ	·		RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM LKALFRTLQAMLLGVWILLLLASLTPLWLYC
]	ł				WRMFPTKGVSP
798 2148	A 6	5528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
			İ	İ	RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
	1	1			LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

SEQ ID NO: of NO: of NO: of NO: of nuclectide sequence   NO: of nuclectide sequence   NO: of nuclectide sequence   NO: of sequence   NO:	ne, ne, n, e  QCGSQSS WRLTKK SVLRIRV ENKLRVV LQMGIE* PMYMGFF MAGIPLLL HDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
nucleotide sequence uenc	QCGSQSS WRLTKK SVLRIRV SNKLRVV LQMGIE* PMYMGFF AGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
eotide sequence  USSN 09/496 09/496 09/496 orrespondi ng to first amino acid residue of peptide sequence  USSN 09/496 09/	QCGSQSS WRLTKK SVLRIRV SNKLRVV LQMGIE* PMYMGFF AGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
sequence    Sequence   1	QCGSQSS WRLTKK SVLRRV SNKLRVV LQMGIE* PMYMGFF AAGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
nence  914  ng to first amino acid residue of peptide residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  915  Peptide sequence  916  Peptide sequence  917  Peptide sequence  918  Peptide sequence  918  Peptide sequence  Peptide sequence  918  Peptide sequence  P	QCGSQSS WRLTKK SVLRRV SNKLRVV LQMGIE* PMYMGFF AAGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
amino acid residue of peptide sequence  amino acid residue of peptide sequence  sequence  T=Threonine, V=Valine, W=Tryptophan Y=Tyrosine, X=Unknown, *=Stop codor /=possible nucleotide deletion, \=possible nucleotide insertion nucleotide insertion  EVFFEGTGLPLPHSDLPTSWCGHSLG FPPAIHENAFIVFIASSLGHMLLTCIL HTVSQEDGLSLAGAPRQPRRKSRTS MVRWELSSNGPNGRGVLGLGLGLG GQNLGL*HCVWVVWETGE*KRWR GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFF REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799  2149  A 6529  1 874  FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	OCGSQSS WRLTKK SVLRIRV SNKLRVV LQMGIE* PMYMGFF AAGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor peptide sequence Y=Tyrosine, X=Unknown, *=Stop codor pepside sequence Y=Tyrosine, X=Unknown, *=Stop codor pepside sequence Y=Tyrosine, X=Unknown, *=Stop codor pepside sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Y=Tyrosine, X=Unknown, *=Stop codor sequence Se	OCGSQSS WRLTKK SVLRIRV SNKLRVV LQMGIE* PMYMGFF AAGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
peptide sequence    Apposible nucleotide deletion, possible nucleotide insertion	QCGSQSS WRLTKK SVLRIRV GNKLRVV LQMGIE* PMYMGFF AGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
requence nucleotide insertion  EVFPEGTGLPLPHSDLPTSWCGHSLG  EVFPEGTGLPLPHSDLPTSWCGHSLG  FPPAIHENAFIVFIASSLGHMLLTCIL  HTVSQEDGLSLAGAPRQPRRKSRTS  MVRWELSSNGNPGRGVLGLGLGLG  GQNLGL*HCVWVVWETGE*KRWRI  GVASRRQ*VRNSVRGLVCHNSSAPF  SPTVFGGGVGG*LHVTFILHPPEVEA  GPSLPQRQGREHIVVILAAPACAPFF  REIRPSP*ELGLRGEPTLSYPASCRVI  RKSYSWKQRLFIINFISFFSALAVYFI  EAGVYTIFAILEYTVVLTNMAFHMT  GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG*  CCLVQGGGDLVDVVQTNHGEDEAG  DEARCKESQQEAQENLREDLCLESF  QIIEGSEREHEETRTKQAALDGEPLG  VHLHPSKEQQGQEGGERQRGARTH  EKGRRVRLRPPSGKLRADQPVRKLG  ELPGLQPHAPTPHTA/PATPTYSPAPF  RWKCPLPVEPRTRQLCRERTRKACI  GLPGDPTGPVTHHAPPVSPTGASGQ  GAVSYAHASATK	QCGSQSS WRLTKK SVLRIRV 6NKLRVV LQMGIE* PMYMGFF AAGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
EVFPEGTGLPLPHSDLPTSWCGHSLC FPPAIHENAFIVFIASSLGHMLLTCIL HTVSQE\DGL\$SLAGAPRQPRRKSRTS MVRWELSSNGNPGRGVLGLGLGLG GQNLGL*HCVWVVWETGE*KRWRI GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFF REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG* CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETTRTKQAALDGEPLG VHLHPSKEQQGGGGGRQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	WRLTKK SVLRIRV ENKLRVV LQMGIE* EMYMGFF AAGIPLLL HDR*WEP IRQPIP*D RHINMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
FPPAIHENAFIVFIASSLGHMLLTCIL HTVSQE\DGLSLAGAPRQPRRKSRTS MVRWELSSNGNPGRGVLGLGLGLG GQNLGL*HCVWVVWETGE*KRWRI GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFE REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETTTKQAALDGEPLG VHLHPSKEQQGQEGGERQGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	WRLTKK SVLRIRV ENKLRVV LQMGIE* EMYMGFF AAGIPLLL HDR*WEP IRQPIP*D RHINMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
HTVSQEUGLSLAGAPRQPRRKSRTS MVRWELSSNGNPGRGVLGLGLGLG GQNLGL*HCVWVVWETGE*KRWRI GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFF REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG*I CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	SVLRIRV SNKLRVV LQMGIE* PMYMGFF AGIPLLL IDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
MVRWELSSNGNPGRGVLGLGLGLGGQNLGL*HCVWVVWETGE*KRWRIGVASRRQ*VRNSVRGLVCHNSSAPFSPTVFGGGVGG*LHVTFILHPPEVEAGPSLPQRQGREHIVVILAAPACAPFHREIRPSP*ELGLRGEPTLSYPASCRVIRKSYSWKQRLFIINFISFFSALAVYFIEAGVYTIFAILEYTVVLTNMAFHMTGNKELLITSQPEEKRFFSALAVYFIEAGVYTIFAILEYTVVLTNMAFHMTGNKELLITSQPEEKRFFSALACDLGGCLVQGGGDLVDVVQTNHGEDEAGDEARCKESQQEAQENLREDLCLESFQIIEGSEREHEETRTKQAALDGEPLGVHLHPSKEQQGQEGGERQRGARTHEKGRVRLRPPSGKLRADQPVRKLGELPGLQPHAPTPHTAPATPTYSPAPHRWCPLPVEPRTRQLCRERTRKACIGLPGDPTGPVTHHAPPVSPTGASGQGAVSYAHASATK	ENKLRVV LQMGIE* PMYMGFF AAGIPLLL HDR*WEP HRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
GQNLGL*HCVWVVWETGE*KRWRI GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFF REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMTI GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG* CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	LQMGIE* PMYMGFF AGIPLLL IDR*WEP IRQPIP*D RHNMYC TAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
GVASRRQ*VRNSVRGLVCHNSSAPF SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFF REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG* CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGGGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	PMYMGFF AAGIPLLL HDR*WEP IRQPIP*D RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
SPTVFGGGVGG*LHVTFILHPPEVEA GPSLPQRQGREHIVVILAAPACAPFE REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	AGIPLLL HDR*WEP IRQPIP*D RHNMYC TAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
GPSLPQRQGREHIVVILAAPACAPFH REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAP RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	IDR*WEP IRQPIP*D RHNMYC TAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	RQPIP*D RHNMYC TAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
REIRPSP*ELGLRGEPTLSYPASCRVI RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTAPATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	RQPIP*D RHNMYC TAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
RKSYSWKQRLFIINFISFFSALAVYFI EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLGY CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPP RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	RHNMYC FAWWDF WCEDWS GGDTDSV FAKDKIL GGGQLTA
EAGVYTIFAILEYTVVLTNMAFHMT GNKELLITSQPEEKRF  799 2149 A 6529 1 874 PFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPP RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	WCEDWS GGDTDSV AKDKIL GGGQLTA
GNKELLITSQPEEKRF  799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	WCEDWS GGDTDSV FAKDKIL GGGQLTA
799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLG' CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	GGDTDSV FAKDKIL GGGQLTA
CCLVQGGGDLVDVVQTNHGEDEAG DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	GGDTDSV FAKDKIL GGGQLTA
DEARCKESQQEAQENLREDLCLESF QIIEGSEREHEETRTKQAALDGEPLG VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLG ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	FAKDKIL GGGQLTA
QIIEGSEREHEETRTKQAALDGEPLO VHLHPSKEQQGQEGGERQRGARTH EKGRRVRLRPPSGKLRADQPVRKLO ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	GGGQLTA
VHLHPSKEQQGQEGGERQRGARTH  EKGRRVRLRPPSGKLRADQPVRKLO ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	
EKGRRVRLRPPSGKLRADQPVRKLO ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	I M W K I T W
ELPGLQPHAPTPHTA/PATPTYSPAPI RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	
RWKCPLPVEPRTRQLCRERTRKACI GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	
GLPGDPTGPVTHHAPPVSPTGASGQ GAVSYAHASATK	
GAVSYAHASATK	
	ERRAEP
800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVI	PGPGGAS
EITFELPDNAKQCFYEDIAQGTKCTI	LEFQVITG
GHYDVDCRLEDPDGKVLYKEMKK	QYDSFTF
TASKNGTYKFCFSNE\FSTFTHKTVY	
ETHLCFLVR/DRVSALTQMESACVS	SIHEALKS
VIDYQTHFRLREAQGRSRAEDLNTR	VAYWSV
GEALILLVVSIGQVFLLKSFFSDKRT	
801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREP	
DGWRRMPRWGLLLLLWGSCTFGLI	
KRIFLKRMPSIRESLKERGVDMARL	
MKRLTLGNTTSSVILTNYMDTQYY	
POTFKVVFDTGSSNVWVPSSKCSRL	
HKLFDASDSSSYKHNGTELTLRYST	
SQDIITVGGITVTQMFGEVTEMPALI	
DGVVGMGFIEQAIGRVTPIFDNIISQ	
VFSFYYNRDSENSQSLGGQIVLGGS	
GNFHYINLIKTGVWQIQMKGVSVG DGCLALVDTGASYISGSTSSIEKLMI	
KRLFDYVVKCNEGPTLPPTFLFLLG	
SADYLFQESYSSKKLSTLAIHAMYI	
\ALGATF\IRKFYTEFDRGNNPHGFA	
802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLE	
LAVVVLLALPVAWGQCNAPEW\LP	
TDEFEFPIGTYLNYECRPGYSGRPFS	
VWTGAKDRCRRKSCRNPPDPVNGN	
	SGDTVIW
DNETPICDRIPCGLPPTITNGDFISTN	RENFHY
GSVVTYRCNPGSGGRKVFELVGEP	SIYCTSND
DQVGIWSGPAPQCIIPNKCTPPNVEI	NGILVSD
NRSLFSLNEVVEFRCOPGFVMKGPI	
LNKWEPELPSCSRVCQPPPDVLHAE	
DNFSPGQEVFYSCEPGYDLRGAASI	
DWSPAAPTCEVKSCDDFMGQLLNC	
NLQLGAK V DF V C DEGFQLKGSSAS	
MESLWNSSVPVCEQIFCPSPPVIPNO	
LEVFPFGKAVNYTCDPHPDRGTSFI	
CTSDPQGNGVWSSPAPRCGILGHC	
FAKLKTQTNASDFPIGTSLKYECRP	LY Y GKPF

NO: of peptide coide sequence uence	000 00	COPA TO	N 4	1 020	The state of	Dagter 3	LA Minoral de Company
muclectide seq- uence  USSN location corresponding of the properties of the period corresponding of the period engage of the period sequence  USSN location of the period of the period of the period sequence period sequence of pepide sequence  Sequence  USSN location of the period of the period sequence period sequence of pepide sequen	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence  USSN			1100	1 .	, ,		
sequence  914  914  ang to first anino acid residue of peptide residue of peptide sequence  915  916  917  918  918  919  919  919  919  919			1	1			
uence    914   ng to first amino acid residue of peptide peptide peptide peptide sequence   Party   Pa			Ì	1			
annino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of per contro	•	uence	ŀ				
residue of peptide peptide peptide peptide peptide sequence    Sequence    Sequence    Sequence    SITCLIDNI, WSSPKDVCKRKSCKTPPDPVNG MYHVITDIQVGSRNYSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRPCGLPPTIANGEDISTINGENSTSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRPCGLPPTIANGEDISTINGENSTSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRPCGLPPTIANGEDISTINGENSTSCTSCHRUPPPV ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPGFVMKUP ENGILVSDNRSLFSLNEVVERCCPF EPPVRIOLUGAKVDFVCDEGFVLKGSSAS GVULAGESLWNSSVPVCEGFTSTDLKVECKP EVYGREPSTCDDLVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDDLVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANGCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYGREPSTCDLDVSSPANCCHCHOOL QAPDHFLAKLKIOTNASDPPIGTSLKYECKP EVYSCEGTSTDLAGASSICTULAGASSIC	uence		ļ	914			
peptide sequence    Popesible nucleotide deletion, "possible nucleotide insertion			ł				
sequence    miclotide insertion			1	1	i e	sequence	
SITCLDNL-WSSPEDVCKRKSCK-TPPPPVNG MYMVTDIQUOSRINNSCTTGHRLIGHSAEGI LSGNAAHWSTKPPICQRPCGLPPTIANGDPIS THREMPITQSVVTYKCPPGSGGGKKYESL-VGE PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV EGIL VSDNRSLFSL-BEVVERCQPPGVMKGP RR VKCQALNK-WFELPSCSS-VCQPPPDVLHA BETQRDEDNISPSGGGEVTSCEGFOSTD-RGAAS MRCTPQGDWSPAAPTCE-VKSCDDFMGQLIA GRVLFPVNLQLGAK-VAYVTCDPHPPGTSSD LIGESTECTSDPQCNGVWSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGILCHC QAPDFLP-AKLKTQTNASDPPGTSLK-VECQP EYYGRPFSITCLDNL-WSSPAPRCGL-PTI ANGDFISTNRENHFYQSVVTTRCNLGSKGKK VEEL VGEPSIYCTSNDDQVGJWSGAPACASCNT- PPFNVENGILVSDMRSLFSLKEV-VEFRCQP GVMKGPRR-VKCQALK-WSPELPSCSR-VCQ PFPELH-GBFITTSHQDVGJWSGAPACASCNT- VEEL VGEPSIYCTSNDDQVGJWSGAPACASCNT- DLGQA-HGRVLFPLNL-QLGAK-VSF-VCDE-GFRL KGSSVSSKCVL-VGMRSGLE-MSV-VSYCDE-GFRL KGSSVSSKCVL-VGMRSGLE-MSV-VSYCDE-GFRL LSVRAGRICKTPQ-PFFASSFTIPNDFEFF-VGTS LYBC-GRYG-WGFSH-STRISTENDFFF-VGTS LYBC-GRYG-GYGWMSST-PRCE LSVRAGRICKTPQ-FFASSFTIPNDFEFF-VGTS LYBC-GRYG-GYGWMSST-PRCE LSVRAGRICKTPQ-FFASSFTINGTVVT-VCQ- TGPDGG-QLFEL-VGGRS-NVCTSNDDQVG-WSS- PFFRCISTNKCTA-PEVENAIR-VY-GRYS-FFS-LITE RFRC-QPF-GYMVGSH-VCQCT-INGR-WGFKL-H- CSRV-CQPPELI-RGEHTL-SR-QDWS-SPS-SI-LEI RFRC-QPF-GYMVGSH-VCQCT-INGR-WGFKL-H- CSRV-CQPPELI-RGEHTL-SR-GDWFS-GGE-VFY SCEPS-TDLEGAS-ALI-CTP-QDWS-E-PRC-TV KSCDDF-LGQL-PRGWVLL-PI-LNL-QLGAKV-SFV-C DGFR-RK-RGR-SS-HC-VLL-VKG-FEI-CTD-QGIWS- QLDH-YCKE-VKC-SFPL-PM-GISKELE-MK-KV-H VGDY-VTL-KCED-GYT-LAG-MS-LA-WY-LP-LI-LI-LI-LS-W- VXTVL/GSSENAL-GVC-SDF-RA-RA-WY-LP-LI-LI-LI-LI-LI-LI-LI-LI-LI-LI-LI-LI-LI-					1 - •		
MYNTYTDIQVGSRINYSCTTOHRLIGHSSAEDL LGSNAAMWSKRPICQRPCCLEPTAINGOFIS THRENFHYGSVVTYRCNPGSGGRKVFELVGE PSIYCTSNDDQVGIWGSPAPQCIIPNCTPPNV ENGILVSDNRSLFSLNEVVERRCQPGFVMKGP RYKCQALNKWEPELPSCSRVCQPPDVLHA ERTQRDDCDNSPGQEVFYSCEPGYDLEGAAS MRCTPGDGWSPAAPTCEWSCSDPHGQLAN GRVLFPVNLOLGAKVDFVCDGFQLKGSSAS YCVLAGMESLWNSSVYVCEQIFCPSPVPVPNQ RHTGKPLEVFPGAAVNYTCDPHDRGTSPD LIGESTRCTSDPQONGVWSSPAPRCGILGHC QAPDHFLFAGLKTGTNASOPPIGTSLKVECRP EFYGRPSITCLDNLWSSSPAPKOELGHCE GAPTHFAGLKTGTNASOPPIGTSLKVECRP EFYGRPSITCLDNLWSSSPAPKOELGHCE EFYGRPSITCLDNLWSSSPAPKOELGHCE FDFVNGMYHVITDQVGSRNYSCTTGHELIG HSSAECILSGNTAHWSTKPPICQRFCGLPPTI ANGDDISTRRENHFYGSVVTYKCH.GSGCRK VEELVGEPSIYCTSNDDQVGIWSGPAPQCIIIPN KCTPPNVENGILUSDNSSLSISLEVCYEPRQQ GFVMKGGPRKVCQALNKWFDELPSCSRVCQ PFPELLHGHTFSHQDNFSSQCEVFYSCEPGY DLRGAASHACTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPIALOLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCRN PALLKGRTTGTPSGDPTYKSESTCTDPHDR GMTFALIGESTIRCTSDPHGNGWSSPAPRCA LSVRAGRICKTPEQPFFASTFIRNDGFFVGTS LNYECRFGYFGKMFSISCLERLWSSVEDNC RKSCGPPEFPNGMVHINITDTGPGSTVNYSC NEGFELIGSRSTTCLVSGNNVTWDKKAPICELI RFRCQPGFVWGSHTVQCOTNGRWGPKLPH RFRCQFGFVWGSHTVQCOTNGRWGPKLPH CSRVCQPFELLIGERTILSQNDNFSSFSISTE IRFRCQPGFVWGSHTVQCOTNGRWGPKLPH CSRVCQPFELLIGERTILSQNDNFSSFGQEVFY SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTPQGDWSPEAPRCTV SCEPSYDLEGAASHCTLAGMKGPKLPH CSRVCQPFELLIGERTILSQNDNFSFSISTE IRFRCQPGFVWGSHTVQCOTNGRWGPKLPH CSRVCQPFELLIGERTILSQNDNFSFSISTE IRFRCQPGFVWGSHTVQCOTNGRWGPKLPH CSRVCQPFELLIGERTILSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLTLSQNDNFSFSISTE IRFRCQDFFLAGRSSHCTLTLAGRSYGPC DEGFFLAGRSSHCTLTLAGRSYGPC DEGFFLAGRSSHCTLTLAGRSYGPC DEGFFLAGRSSHCTLTLAGRSYGPC DEGFFLAGRSSHCTLTLAGRSYGPC DEGFFLAGRSSHCTLTLAGRSYGPC DEGGRAFTE VARAMFAGRAFTICTTSGSPWGCQAADDRWD PPLAKCTSRTIDALITGSFSPAFCAURCH				<u> </u>	sequence		
LISGNAAHWISTKPPICQRIPCGLIPTIANGDPIS TNRENFPIGYSVTYSCPPOSGGRKYKELVGE PSIYCTSNDDQVGIWSGPAPCCIIPNKCTPPNV ENGIL VSDNRSLFSLBEVVERRCQPPFDVILANG RIVKCQALINK WEPELPSCSRVCQPPPDVILAN ERTQRDKDNFSSOQEVFYSCEPGYDLRGAAS MRCTPQGDWSPAPTCEVKSCDDFMQQLIN GRVLFPVNLQLOAKVDFVCDGFCLKGSSAS YCVLAGMESL WINSSYVPCUGICFSPPVIRNG RHTGKPLEVFFGKAVNYTCDPHPDRGTSFD LIGESTRETSDPQGNGWSSPAPRCGUGHC QAPDHFLFAKLKTQTNASDPFIGTSLKYECRP EYYGRPFSTICLDNLWSSRAPVCKSKSCKTPPDYNGMYHYTDQVGSKINYSCTTGHRLIG HSAAGLISGNTAHWSTKPPICQRPGGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNDDQVGIWSGAPACGLPTI ANGDPISTNRENFHYGSVVTYRCNLGSRGKK VEEL VGEPSIVCTSNGWSFAPACGL LOQLPHGRVLFPLNLQLGAKVSFVCDEGFEL LGQLPHGRVLFPLNLQLGAKVSFVCDEGFEL LGQLPHGRVFFLNLGGSFSTAPCT RGMSTALLGESTIRCTSDPHONGVWSSPAPACCL LSVRAGHCXTPGOPFASSTTPNDFEFFV GTS LNYECRPOYFGKMFSISCLELLW WSSVEDNC RKSCGPPPEPFRMYMTHINTDTGFSTVNYSC NEGFELLGSRSTNCTSKDDQVGVWSSS PPRCISTNRCTAPEVENARVPGMSFSTSLTEL RFRCQPGFVERMRFSTSLTEL RFRCQPGFVERMRFSTSLTEL RFRCQPGFVERMRFSTSLTEL RFRCQPGFVERMRFSTSLTEL RFRCQPGFVERMRFSTSLTEL RFRCQPGFVLLVGKGFTCTDQGIWS SPAPACELPVGAACPHPFLNDGHTVGGWSSVTP COPPTBLINGGTTLSQDWSPCAACHT VGGPVTLLFLNLQLGAKVSFVC DEGFFLLKGRSASHCVLAGMKALWNSSVPVC COFFCTPPALINGARTUTGLDGIPVGKFSVST COPHERGMTFLLIGESTIRCTSEPHGNGWS SPAPACELPVGAACPHPFLNDGHTVGGHVSL VLPGMTTSLTCDGGVTLFUNLQLGAKVSFVC DEGFFLLKGRSASHCVLAGMKALWNSSVPVC COFFTPPALINGARTUTGLGDTVGGHVSUSSTAPC COPPTALINGARTUTGLDGIPVGGBVST ALKERKONNAHDPKEVAHLLILLELSWI LKHRKGDNAHCDPKEVAHLLILLELSWI LKHRKGDNAHCDPKEVAHLLILLELSWI LKHRKGDNAHCDPKEVAHLLILLELSWI LKHRKGDNAHCDPKEVAHLLILLL LLLPFLLLLLRGSHAGCNLTVAAVVLPLANTSY VKSTLGSSENALGVCGAAALITTKYDGEBHCFT VAGLARGCAAALITYATCDGEBHCFT VAGLARGCAAALITYATCDGEBHCFT VAGLARGCA							
TINEENFHYGSVYTYRCNPGSGGRKVFELVG BPISYCTSNDDQVGWSGPAPQCIPPNCTPPNV ENGIL VSDNRSLFSLNEVVFRCQPGFVMLGA RRVKCQALNKWPFELPSCRYCQPFPVMLGA RRVLCQALNKWPFELPSCRYCQPFPVMLGA RRVCQALNKWPFELPSCRYCQPFPVMLGA AMRCTPGGWSPAPTCEVKSCDFMGQLAS MRCTPGGWSPAPTCEVKSCDFMGQLAS MRCTPGGWSPAPTCEVKSCDFMGQLAS AMRCTPGGWSPAPTCEVKSCDFMGQLAS YCVLAGMESLWANSSVPVCEDGFGLKGSSAS YCVLAGMESLWANSSVPVCEDGFGLKGSSAS YCVLAGMESLWANSSVPVCEDGFTSTBLT LIGESTRCTSDPQGNGVWSSPAPRCGILGHC QAPDHFLFAKLKTQTINASDPFIGTSLKYECRP EYYGRFFSITCLDILVWSSPKDVCKRKSCKT PPPVNGMYHVIDIDQ VGSRWSYSCTIGHRLIG HSSABCLLSGNTAHWSTKPFICQRTCGLFFTI ANGDFTSTNRENHYGSVYTYRCNLGSRGRK VFELUGEPSIYCTSNDDQVGIWSGRAPQCILIH KCTPPNVENGILVSSNRSLFSLNEVVERCQC GFVMKGPRVKCQALNKWEPELPSCSRVCQ GFVMKGPRVKCQALNKWEPELPSCSRVCQ PPPELHAGHTPSSNNSLFSLNEVSFVCFGCPQ GFVMKGPRVKCQALNKWEPELPSCSRVCQ PPPELHAGHTPSTDDPYGGRSSYTCOPHPDR GMTSLLGGSTRTCTSDPHGNGVWSSPAPRCE LGURAASLHCTPQGDWSPSAPRCAVKSCDDF GMTSLLGGSTRTCTSDPHGNGVWSSPAPRCE LSVRAGHCKTEGPFPASPTIFINTGPYGTSVTNYC RRSSCGPPPEPFNGTPTSDDPYGGRSYTCDPHPDR GMTFNLIGESTRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTEGPFPASPTIFINTGPTYTYCQCH RRSSCGPPEPFINTGPTSSNNTSTENTGTVYTYCQCH RGSSCSTRTCLVSGNNVTWDKKAPICIC SCEPPFTISNGGPYSNNTSTENTGTVYTYCQCH TGPDGGQLFELVGERSIYCTSCDDQVGVWSS PPPRCISTNKCTAFEVENARWPGNSFFSLTEI IRFRCQPGFFWWGSHTVCQCTNGRWFFSLTEI IRFRCQPGFFWWGSHTVCQCTNGRWFFSLTEI IRFRCQPGFFWWGSHTVCQCTNGRWFFSLTEI IRFRCQPFPPANGHTNTLIGESTTRTSEPHGNGVWS SPAPRCELPUGAACAPHPSVLNGNFSSSLTEI IRFRCQPFPPANGHTNTGHTGPTGDFTASPTSCHOWWS SPAPRCELPUGAACAPHPSVLNGNLGAKWSFVC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWNSSVYC DEGFRLKGRSASHCVLAGMALWALWNSSVYC DEGFRLKGRSASHCVLAGMALWALWNSSVYC DEGFRLKGRSASHCVLAGMALWALWALWALWH TUTTLIGGTTTRLLIFLL TAGAALAGTATAWALWALL TAGAALAGTATAWALPLLAGMALWALWALL TAGAALAGTATAWAL							1
PSINCTSINDQVGIWSGPAPQCIIPNICTPPNV ENGIL VSDINSLI-SINEVEFEROFEWYSCOPPPOVIMGE RRVKCQALNKWEPELPSCSRVCQPPPOVIGAAS MRCTPOGDWSPAAPTCEVKSCODENGQLIN GRVLEPVNLQLGAKVDFVCDGFCLKGSSAS YCVLAGMESLWMSSVPVCEOFECFSPPVIPNO RHTGKPLEVFPGAAVNYTCDFHDROTSIS- LIGESTIRCTSDPGONGVWSSPAPCCILI-GHD QAPDHFLFAKLKTQTNASDFHGTSLKYECG- EY GRFFSITCLIN LWSSYRDVCKKSSCKTP PDFVNGMYHVITDIQVGSRINYSCTTGHEIGHD GASEACLLSGNTAHWSTKPPICQRPCGLPPT ANGDFISTINENFHYGSVVTYRCRLGSRGK VFELVGEPSYCTSINDDQVGIWSGPAPQCIIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQC GFVMKGPRVKCQALINKWEPELPSCSRVCQ PFPELLHGEHTTSHQDNFSRGQEVFYSCEGF DLRGAASHCHTPQDWSPSAPRCAVKSCDDP LGQLPHGRVLFPLIN_QLGAKVSFVCDEGFT KGSSVHCVL VGMRSLWNNSVYCEHIFCRN PALLNGRHTGTSDPHONGVWSPAPRCE LSVRAGHCKTEGOPPASPTIPNDEEPPVGIN KGSSVHCVL VGMRSLWNNSVYCDHIFCRN PFALLNGRHTGTSDPHONGVWSPAPRCE LSVRAGHCKTEGOPPASPTIPNDEEPPVGIN KGSSVHCVL VGMRSLWNNSVYCDHIFCRN PFALLNGRHTGTSDPHONGVWSPAPRCE LSVRAGHCKTEGOPPASPTIPNDEEPPVGIN KGSSVHCVL VGMRSLWNNSVYCDHIFCRN PFALLNGRHTGTSDPHONGVWSPAPRCE LSVRAGHCKTEGOPPASPTIPNDEEPPVGIN KGSSVHCVL VGMRSLWNSVYCDHIFCRN PFALLNGRHTGTSDPHONGVWSPAPRCE LSVRAGHCKTEGOPPASPTIPNDEEPPVGIN KGSSPFTISNODPVGNRSFFSLTEI IRFRCQPGFFWNVSHSISLELBU VWSSVEDDIS LNYEGERSTYLCVGRNSTFSLTEI IRFRCQPGFFWNVGSHTVQCQTINGRWGFKLV SCEPPYPLROADSHAWSVEDNC VGCPPPFLINGRFTLIHGFTLYTQCDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTAGGBSSVUP VSCCPPSTSTAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTQGDWSPEAPRCT VSCCPPPPBLAGAASHCTTAGACPSVALULUT VSCCPPPPBLAGAASHCTTAGACPTVLLLLLL TAGAALGFGVCAAAVGCAGTAAFLAGAVCLUP VALTURENSGWCAAAVGCAAATAGAAAVCLLAGAACHTAGAAAACAAAACAACHTAGAAACAAACAACAACAACAACAACAACAACAACAACAAC							LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS
ENGIL VSDNRSLESLNEV VEFRCQOFPY VLHA ERTQRDKDNFSPGGEVFYSCEPGYDLRG AAS MRCTPGGBWSPAPTCEVKSCDPHOVLHA ERTQRDKDNFSPGGEVFYSCEPGYDLRG AAS MRCTPGGBWSPAPTCEVKSCDPHOVLHA GRVLFPYNLQLGAKVDFVCDGIFCSFPVING RITGRELEVFFFCAK ANVITCDFHDRGTSFD LIGESTIRKTSDPQGNGVWSSPAPRCGILGHO QAPDHELFAKLKTUTNASDPFISTSLKYECRP EYYGRFSTICLDNLVWSSPKDVCKRKSCKTP PPDFVNGMYHVIDDIOVGSRNYSPAPRCGILGHO HSSABCILSGNTAHWSTKPFICQRIPGLFPTI ANGDFISTNRENFHYGSVVTTKCNLGSRGRY KCTPPNVENGILVSDNSLSISLNEVVEFRCQP GFVMKGPRRVKCQALNK WEPELPSCSRVC) GFVMKGPRRVKCQALNK WEPELPSCSRVC) PPELLIGBHTPSINDDNFSPGGEVFYSCEPGY DLRGAASIHCTPQDDWSPSAPRCAVKSCDDF LQQLPHGRVLFILNLQLGAKVSFVCDEHGCN PPELLIGBHTPSICDDNFSPGGEVTYSCEPGY DLRGAASIHCTPQDDWSPSAPRCAVKSCDDF GMTMLIGESTIRKTSDPHGNOV WSSPAPRCE LSVRAGHCKTFEGFFASPTIENDEFFVCTS  RKSSSSHCVLVGMRSLWNNSVEVCEHFCCN PPALLNORHTGTPSGDIPYGKEISYTCDPHDP GMTFNLIGESTIRCTSDPHGNOV WSSPAPRCE LSVRAGHCKTFEGFFASPTIENDEFFVCTS NEGFFLIGSFSTTCLVSGNNVTVMSVEDDNC RKKSCGPPPFSNNRTSFHNGTVMTYQCH TGPDGGQLFELVGRSINCTSDDQGVWSS PPRRCISTNKCTAFEVENAHRYPGNSFSFSLIL ERFCGPGFFWMGSTSLCELNLVWSSVEDDNC RKSCGPPPFINGFYSNNRTSFHNGTVMTYQCH TGPDGGQLFELVGRSINCTSDDQGVWSS PPPRCISTNKCTAFEVENAHRYPGNSFSFSLIL ERFCGPGFFWMGSTHTVOLTGNRWGFKLFH CSRVCQPPPELHGEHTLSHQDNFSPGGEVTY SCEPSYDLRGAASHLCTPGGDWSPAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFFLKKRASSRCVLAGKKALWNSSVEDV SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVYS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVYS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTIRRTSEPHGNGVS SPAPRCELPVGAACHTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTRATSEPHGGSSLLLLL KVYLL SGGDTALAGRACHTTPTGLDBYGKEVSYT CDPHDRGMTFNLIGESTRATSEPHGGSSLLAKKVYH YGDYVTLKCEDCYTLAGSFWSQQAADBYGGSSVHP TLATTMPRKGKYVTYCSSPDARTHAWAYPLL TAGAPALGFGVKAAPVOLTFLAITTS PWSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSSEBALGOVSADAGHAAVTAVPLLATTS PWSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSGSBALGOVSADAGHAAVTHYPLLL TAGAPALGFGVKAAPVGGEBLCFC LV							TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
RRYKCQALNKWEPELPSCSRVCQPPPDYLGAAS BRCTPQGDWSPAAPTCEVKSCDDFMGQLIA BERTQRDKDNFSPGQEVFYSCEPGYLRGAAS MRCTPQGDWSPAAPTCEVKSCDDFMGQLIA GRVLPVALOLGAKVDPVCDGFGQLKGSSAS YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG RHTGKPLEVFFGKAVNYTCDFHDRGTSTS- LIGESTIRCTSDPQGNGVWSSSPAPCGLIGHC QAPDHFLFAKLKTQTNASDPFIGTSLKYECGF EYYGRFSTICLDLIAVWSSKPVCKKRSCKTP PDFVNGMWHTDIDQVGSRNYSCTTGHRLIG HSSASCLISGNTAHWSTKPPICQRPGCIPPT ANGDFISTINENPHYGSVVTYRCRLGSRGK VFELVGEPSYCTSINDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELOFERQQP GFVMKGPRRVKCQALNKWEPELOFERQPO GFVMKGPRRVKCQALNKWEPELOFERQPO GFVMKGPRRVKCQALNKWEPELOFERQPO GFVMKGPRRVKCQALNKWEPELOFERQPO DLRGAASLHCTPQGDWSSPAPRCAVKSCDDF LGQLHGRAVIFIN,QLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCFN PPALLNORHIGTPSGDPYGKSEYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAAGHCKTEGPFPASPTIPINDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGPFASPTIPINDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGWSSPAPRCE SCEPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGWSTNYCC NEGFRLIGSSTTCLVSGNNVTWDKKAPICLE SCEPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGWSTNYCC NEGFRLIGSSTTCLVSGNNVTWDKKAPICLE SCEPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPFINGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPEPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSISCLENLWSSVEEDOL RKSCGPPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQYFGKMFSTSLTCH RKRCGPTMANGSTSHTTLIJGSTGSTNTYCC RKSCGPPPTISNGGFYSNNTTSHENDFEPPVGTS LNYECRFQFYMKGSSNTTYTCLUSHTS PPRCELPUFGATTTNIGGSTURTTGGFYSALCUS LNYECRFQFYMTGGFYSTALUS LKHRKGMNAHENFEEVAHLHRSQGGSSVHP RKLTTTHLIILIFLUSH LKHRKGMNAHENFEEVAHLHRSQGGSSVHP RKLTTTHLIILIFLUSH LKHRKGMNAHENFEEVAHLHRSQGGSSVHP RKLTTTHLIILIFLUSH LKHRKGMAHAUCHSLTTHTAGFYAKLGUEDOL							
ERTQRDKDNFSPGGE/FYSCEPGYDLRGQLIN GRVLFPVNLQLGAKVDFVCDEGFQLKGSAS, MRCTPGGDWSPAPTCEVKSCDFMGQLIN GRVLFPVNLQLGAKVDFVCDEGFQLKGSAS, YCVLAGMESL WNSSVPVCDEGFQLKGSAS, YCVLAGMESL WNSSVPVCDEGFGLKGSAS, YCVLAGMESL WNSSVPVCDEGFGLKGSAS, YCVLAGMESL WNSSVPVCDEGFGLKGSAS, YCVLAGMESL WNSSVPVCDEGFSTSD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPDHFLFAKLKTQTNASDPFJGTSTLKYECRP EYYGRFSTICLDNLVWSSPKDVCKRKSKFT PDFVNGMYWITDIDQGSRNTYSCTTGHRLIG HSSABCLLSGNTAHWSTPPICQRTCGLFFTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEFSYTCTSNDDQVGIWSGRAPQCILIN KCTPFNVENGIL SSDNRSLFSLINEVVEFRCQP GFVMKGPRVKCQALNKWEFELFSCSRVCQ PFELLIGBHTFSIDNNSLSGLEVFRCQP GFVMKGPRVKCQALNKWEFELFSCSRVCQ PFELLIGBHTFSIDNONFSNGGEVTYSCEFGQ PPMELIGBHTFSIDNODFSNGGEVTYSCEFGQ DLRGAASHCTPQODWSPBAPRCAVKSCDDF LQQLPHGRUFLPLQ QLGAKVSCPDGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCRN PALLNGRHGTGTSGDPYGKEISYTCDPHDPA GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTFEQFFASFTIFINDEFFYCTS  RKKSCGPFPFSNGWFHINTDTCFGSTVNTYS CNEGFFYSINGFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGRSINCTSKDDQVGVWSS PPPRCISTNKCTAFEVENAIRVPGNRSFSSLIL ERKCQGFFYMSGNTYCQCTMGWGPKLFH CSRVCQPPFBLHGGHTLSHQDNFSPGGEVTY SCEPSYDLRGAASHCTPGGDWSPEAPRCTY KSCDDPLGQLFHCWLLDFALQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC DEGFRLKGRSASHCVLAGMCANTTVHLILIBLSW HINTAGRAGACALTTATAGTVTHLLILLL TAGAPALOFGERCHTLGGGGSVHP TLLTTHENGERVLP TLATTHTTHAGKGRVTYCTCSPD			ł				
MRCTPQGDWSPAAPTCEVKSCDDFMQGLIN GRVLFPVIN_LOLAGKVDFVCDEGFQLKGSSAS YCVLAGMESLWNSSVPCEQIFCCFSPPVIPNG RHTGKPLEVFPGKAVNYTCDPPDRGTSFD LIGESTIRCTSDPQONGVWSSPARCGILGHC QAPDHFLFAKLKTQTNASDPJGTSLKYCCRP EYYGRPSTICLDNILWSSSFNDVCKRKSCKTP PDPVNGMVHVITDIQVGSRNYSCTGHRLIG HSSAECLSGNTAHWSTKPJCQRIPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDQVGJRAPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDQVGJRAPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDQVGJRAPCCIPC GFVMKGFRRAVCQALAKWEPELPSCSRVQ QFPFELHGEHTPSHQDDNFSPQGEVFYSCEGGY DLRGAASLHCTPQGDWSPEAPRCAVSCODF LGQLPHGRVLFPLALQLGAKVSPVCDEGFRL KGSSVSHCVLVGMRSLWNSVPVCEHFCND PALLINGRHTGTPSGDPTVGREISYTCDPHDPR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPGOFFASTIPINDFEFPVGTS LNYECRGGYFGKMFSISCLENLVWSVEDNC RRKSCGPPFEFFNGMVHINDTQFGSTVNYSC NEGFRLIGSPSTTGLVSGNNVTWDKKAPICEII SCEPPPTISNGDPSSNNTUTTDTQFGSTVNYSC NEGFRLIGSPSTTGLVSGNNVTWDKKAPICEII SCEPPPTISNGDPSSNNTUTTDTQFGSTVNYSC NEGFRLIGSPSTTGLVSGNNVTWDKKAPICEII SCEPPPTISNGDPSSNNTUTTDTQFGSTVNYSC NEGFRLIGSPSTTGLVSGNNVTWDKKAPICEII SCEPPPTISNGDVSSNNTVTWDKAPICEII SCEPPPTISNGQVSSVC NEGFRLIGSPSTTGLVSGNNVTWDKAVSVVC USCHOPPELINGEHTLSHQDNNSPGQVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPIALQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQFFCNPPALINGRHTGTPLODIPYGKEVSTT CDPHPDRGMTFILIGESTRRETSPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHTIGGHVSL YLPGMTISTTCDGYTLLVGKGFFCTDQGINS QLDHYCKEVNCSPPLFAMNGIKELEMKXVSPVC EQFFCNPPALINGRHTGTPLODIPYGKEVSTT CDPHPDRGMTFILIGESTRRETSPHGNGVWS SPAPRCELPVGAACPHPRIQNGHTIGHVSL YLPGMTISTTCDGYTLLVGKGFFCTDQGINS QLDHYCKEVNCSPPLFAMNGIKELEMKXVSPVC EQFFCNPPALINGRHTGTPLODIPYGKGVSSTV BTLTTTTCDGYTLLVGKGFFCTDQGINS QLDHYCKEVNCSPPLFAMNGIKELEMKXVSPVC EQFFCNPPALINGRHTGTPLODIPYGKACH UKHKKGNNAHENFREVAHILHSQGGSSVHP RTLQTNEENSRVLP HGRSARLAAVPAEAMFGPRRPRGSGSSVHP RTLQTNEENSRVLP HGRSARLAAVPAEAMFGPRRPRGEGHCFF LVGGLFMRVRGRALLMLYAVPLLENTSY PWSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVGSDTAAPLAAVLLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGGGGGAGAPAFTAMLLLAGLHTYADPNSYLV LTAGAPAPREW ERGDGGDVSARAQAFQAAKHTYKDPNDFSYL BEGGGGDGDVSARAQAFQAAKHTYKDPN			ł	}			RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS YCVLAGMESLWSSVPVCDFSPPVPING RHTGKPLEVFPFGRA VNYTCDFHEDRGTSFD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPDHELFAKLKTQTNASDPFIGTSLKYECRP EYYGRFSITCLDNL WSSKDVCKRKSCKTP PDPVNGMVHVITDQVGSRNYSCTTGHRLIG HSSAECILSGNTAHWSTKPFICQRECGLPTI ANGDFISTNRENFHYGSVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIJPN KCTPPVERGIL VSDRSLFSEVVERCOLPTI ANGDFISTNRENFHYGSVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIJPN KCTPPVERGIL VSDRSLFSEVVERCOLPTI ANGDFISTNRENFHYGSVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIJPN KCTPPVERGIL VSDRSLFSEVVERCOLPTSCHOOL GFVMKGFRRVKCQALAKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPQGEVFY3CEPGY LIGQLPIGHTVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMSL WNNSVPVCEHFCFN PPALLNGRHTGTPSGDPYGREISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQPFPASPTIPNDFEFFVGTS LNYECRRGVYGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFRAGMTWINDTDTGFGSTVNYSC NEGFRLIGSFSTTCLVSGNNVTWDKCKARICELI SCEPPPTISNGVKMSSISCLENLVWSSVEDNC RRKSCGPPPEPFRAGMTWINDTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT			ĺ			•	ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS YCVLAGMESLWSSVPVCDFSPPVPING RHTGKPLEVFPFGRA VNYTCDFHEDRGTSFD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPDHELFAKLKTQTNASDPFIGTSLKYECRP EYYGRFSITCLDNL WSSKDVCKRKSCKTP PDPVNGMVHVITDQVGGRINYSCTTGHRLIG HSSAECILSGNTAHWSTKPFICQRECGLPTI ANGDFISTNRENFHYGSVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIJPN KCTPPVERGIL VSDRSLFSEVEFRCOP GFVMKGRRRVKCQALAKWEPELPSCSRVCQ PPFELLHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLICTPQGDWSEAPRCAVKSCDDF LGQLPIGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMSL WNNSVPVCEHFCFN PPALLNGRHTGTPSGDDPYGREISYTCDPHDR GMTFNLIGESTRICTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQPFASPTIPINDEFFVGTS LNYECRRGVYGKMFSISCLEIN LWSSVEDNC RRKSCGPPPEPFRAGNTWNTDVCKARICEIN SCEPPFTISNGVKMFSISCLEIN VWSSVEDNC RRKSCGPPPEPFRAGNTWNTDVCKARICEIN SCEPPFTISNGVKMFSISCLEIN VWSSVEDNC RRKSCGPPPEPRAGNTWNTDVCKARICEIN SCEPPFTISNGVKMFSISCLEIN VWSSVEDNC RRKSCGPPPEPRAGNTWNTDVCKARICEIN SCEPPFTISNGVKMFSISCLEIN VWSSVEDNC RRKSCGPPPEPRAGNTWNTDVCKARICEIN SCEPPFTISNGVMSSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGKLPH CSRVCQPPFELHGGEHTSLAPUFONSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGKLPH CSRVCQPPFELHGGHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAOMKALWNSSVPVC EQFFCNPPALINGRHTGTPLOGIPYGKEVSTT CPHPPDRGMTFNLIGESTIRTSEPHGNGVWS SPAPRCELPVGAACPHPFKIQNGHYGGHVSL YLPGMTSSTTCDFGYLLVGKGFFFCTDQGIWS CDAFYCKEVNCSPLFMGISKELEKKVYH YGDYVTLKCEGGVTLEGSPACQAADDRWD PPLAKCTSRTDAAJVCKWGSPCAADDRWD PPLAKCTSRTDAAJVCAVGFFCTDQGIWS CLDHYCKEVNCSPLFMGISKELEKKVYH YGDYVTLKCEGGVTLEGSPCQAADDRWD PPLAKCTSRTDAAJVCAVVLPLANTYVPLL LILHPLLLILRGSHAGNLTVAVVLPLANTYPLL TAGAPALGGVKDCAALAQVCARPCLLFGWT RILKRTMPRKGRAILGVSDTAAAVVLFLLANTYPLL TAGAPALGGVCDEYALTTAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPODEEHCFF LVGGLFMRVRDRINTTVDHLEFAEDDLSHYT RLLKTMPRKGRAJMLYAYAPRADEHCFF LVGGLFMRVRDRINTTVDHLEFAEDDLSHYT RLLKTMPRKGRAJMLYAYAPROBEHCFF LVGGLFMRVRDRINTTVDHLEFAEDDLSHYT RLLKTMPRKGRAVINCSSPAFRTMLLLALGLEICH LLUPPLLLEICRGAGAGAGPAYAKLDGF				1			MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
RHTGKPLEVFFPGKAVNYTCDPHPDRGTSFD LIGESTIRCTSDPQGNGWVSPRPACGILGHC QAPDHFLFAKLKTQTNASDPPIGTSLXYECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVTDLQVGSRINYSCTTGHRLIG HSSAECLLSGNTAHWSTKPPICQRIPCGLPTI ANGDFISTINENHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALMKWEPELPSCSRVCQ PPPELLIGEHTISHQDNFSDQGVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHIGRVIFPLNLQLGAKVSPVCDEGFRL KGSSVSHCVLVGMRSL WNNSVPVCEHFCPN PPALLNGRHTGTPSGDDPYGKEISYTCDPHDR GMTNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQPFFASPTIPINDPEFFVGTS LNYECRRGYGKMFSISCLEINLVWSSVEDNC RRSSCGPPPEFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGFYSNNTSFHNGTVVTTVQCH TGPDGBQLFELVGERSIYCTSKDDQVGWVSS PPPRCISTNKCTAPEVENAIRVFORNSFFSLTEII RFRCQPGFVMVGSHTVQCQTNGRWGKLPH CSRVCQPPPELINGEHTISHQNNFSQGEVFY SCEPSYDLRGAASLHCTPQCDWSPEAPRCTV KSCDDFLGGLPERGNLIFTLNQLGAKVSFVC DEGFRLKGRASSHCVLAGMKALWNSSVPVC EQIFCPNPALLNGRHTGTFLLODIPYGKEVSYT CDPHPDRGMTFNLIGESTIRKTSEPHGNGWWS SPAPRCELPVGAACPHPPKIONGHYIGGHVSL YLPGMTISYTCDPGYLL VGKGFFFCTDQGIWS QLDHYCKEVNCSPPLFMNGISKELEMKVYH YGDVYTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTEDALIVGTLSGTTEFILLIELSWI ILKHRKGNNAHENPKEVAHLHSQGGSSVHP RTILQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMFGPRRPAGSRLRLLLL LILPPLLLLIRGSHAGHLTVAVVVLPLANTYPLL TAGAPALGFGVLDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEBHCTF VRSWSMARVGPAVELALAQVKARPDLLFGWT VRTVLGSSENALGVCSDTAAAVVCRTTAATWRVPLL TAGAPALGFGVLDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEBHCTF LVGGLFMRVRDRLNITTVDHLEFAEDDLSHYT RILLRTMPRKGRVINCSSPDAFARTMLLALLEA LLTMPRKGRVINCSSPDAFARTMLLALLEA LLTMPRKGRVINCSSPDAFARTMLLALLEA GLCGEDYVFFHLDIFFGQSLQGGGGAPARRPW ERGDGGDVSAAAAPQGAFTAAMLALLEA ERGDGGDDVSARAQAFQAAALTTYNDPDTPSYL RILLTMPRKGRAVHLTSCDPDTSPLAFRIPW ERGDGGDGDVSARAQAFQAARTTAMLALEA GLCGEDYVFFHLDIFFGQSLQGGGGAPRRPW				ĺ			
RHTGKPLEVFFPGKAVNYTCDPHPDRGTSFD LIGESTIRCTSDPQGNGWVSPRPACGILGHC QAPDHFLFAKLKTQTNASDPPIGTSLXYECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVTDLQVGSRINYSCTTGHRLIG HSSAECLLSGNTAHWSTKPPICQRIPCGLPTI ANGDFISTINENHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALMKWEPELPSCSRVCQ PPPELLIGEHTISHQDNFSDQGVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHIGRVIFPLNLQLGAKVSPVCDEGFRL KGSSVSHCVLVGMRSL WNNSVPVCEHFCPN PPALLNGRHTGTPSGDDPYGKEISYTCDPHDR GMTNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQPFFASPTIPINDPEFFVGTS LNYECRRGYGKMFSISCLEINLVWSSVEDNC RRSSCGPPPEFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGFYSNNTSFHNGTVVTTVQCH TGPDGBQLFELVGERSIYCTSKDDQVGWVSS PPPRCISTNKCTAPEVENAIRVFORNSFFSLTEII RFRCQPGFVMVGSHTVQCQTNGRWGKLPH CSRVCQPPPELINGEHTISHQNNFSQGEVFY SCEPSYDLRGAASLHCTPQCDWSPEAPRCTV KSCDDFLGGLPERGNLIFTLNQLGAKVSFVC DEGFRLKGRASSHCVLAGMKALWNSSVPVC EQIFCPNPALLNGRHTGTFLLODIPYGKEVSYT CDPHPDRGMTFNLIGESTIRKTSEPHGNGWWS SPAPRCELPVGAACPHPPKIONGHYIGGHVSL YLPGMTISYTCDPGYLL VGKGFFFCTDQGIWS QLDHYCKEVNCSPPLFMNGISKELEMKVYH YGDVYTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTEDALIVGTLSGTTEFILLIELSWI ILKHRKGNNAHENPKEVAHLHSQGGSSVHP RTILQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMFGPRRPAGSRLRLLLL LILPPLLLLIRGSHAGHLTVAVVVLPLANTYPLL TAGAPALGFGVLDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEBHCTF VRSWSMARVGPAVELALAQVKARPDLLFGWT VRTVLGSSENALGVCSDTAAAVVCRTTAATWRVPLL TAGAPALGFGVLDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEBHCTF LVGGLFMRVRDRLNITTVDHLEFAEDDLSHYT RILLRTMPRKGRVINCSSPDAFARTMLLALLEA LLTMPRKGRVINCSSPDAFARTMLLALLEA LLTMPRKGRVINCSSPDAFARTMLLALLEA GLCGEDYVFFHLDIFFGQSLQGGGGAPARRPW ERGDGGDVSAAAAPQGAFTAAMLALLEA ERGDGGDDVSARAQAFQAAALTTYNDPDTPSYL RILLTMPRKGRAVHLTSCDPDTSPLAFRIPW ERGDGGDGDVSARAQAFQAARTTAMLALEA GLCGEDYVFFHLDIFFGQSLQGGGGAPRRPW							YCVLAGMESLWNSSVPVCEOIFCPSPPVIPNG
LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPHHELPALIKITOTIASPIGTSLKYECCP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVITIDIQVGSRRYSCTTGHRLIG HSSAECLLSGNTAHWSTRPICCRIPCGLPTI ANGDPISTINRENFHYGSVVTYRCINGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNSEVPERCQP GFVMKGPRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVSCODDF LGQLPHGRVLFUNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHFCCPN PPALINGRITGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTEOFPPA SITVNDSKAPRCE LSVRAGHCKTEOFPPA SITVNDSKAPKCE LSVRAGHCKTEOFPPA SITVNDKKAPICEII SCEPPFIISNGDFYSNNRTSFHNGTSTNNYSC NEGFRLIGSSFTICLVSGNNVTWDKKAPICEII SCEPPFIISNGDFYSNNRTSFHNGTSTNYSC NEGFRLIGSSFTICLVSGNNVTWDKKAPICEII BFRCOGEFWOSSHTVCTSKDDQVGVWS PPPRCISTNKCTAPEVENARVPGNNSFFSLTEI BFRCOGEFWOSSHTVCTSKDDQVGVWS PPPRCISTNKCTAPEVENARVPGNNSFFSLTEI BFRCOGEFMVGSHTVCTYCCH CSRVCQPPEILHGEHTLSRQDNFSPGCEVFY SCEPSYDLRGAASLHCTPQGDWSPFPARCTV KSCDDFLGQLPHGRVLLPLINLQLGAKVSFVC DGGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCNPPPALINGRHTGLDPFYGREVSYT CDPIPDRGMTFNLIGESTIRRTSEPPIGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLFGMTISYTCDFGYLLVCKGFIFCTDQGIWS QLDHYCKEVNCSPPLFMNGISKELEMKKVYH YGDYYTIK.CEGGYTLGSFWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTTFLLIIELSWI LKHRKGNNAHENFREVAHLHISQGSSVHP RTLOTNEENSRVLP HGRSARLAAVPAEAMPGPRRPAGSSSVHP RTLOTNEENSRVLP BGRSARLAAVSFALALAAVOLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRYPLL TAGAPALLGFGVKDFVALTTRAGFSYAKLLOF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVGGLFWRVEDRINITVDHLEFEAEDDLSHYT RLRTMFRKGRVYICSSPAFRILMLLALALE GLCGEDYYFFHLDIFGQSLQGGGGPAPRRPW ERGGGDQDVSARQAFPARALILALALE BCGCCGDYYFFHLDIFGGSLQGGGFPARRPW BERGGGGDVSARQAFARALILALALALE BCGCCGDYYFFHLDIFGGSLQGGGFPARRPW BERGGGGDVSARQAFFQAAKIITKKOPPDNFEYL BERGGGGDVSARQAFFGAAKIITKKOPPDNFEYL BERGGGGDVSARQAFFGAAKIITKKOPPDNFEYL BERGGGGGARPARPW BERGGGGGDVSARQAFFGAAKIITKAGPSPAARLLALE BCGCGGDYFFHUDIFGGSLQGGGGFPARRPW BERGGGGDVSARQAFFGAAKIITKKDPPDNFEYL BERGGGGGGARPARPW BERGGGGGGGARPARPW BERGGGGGGGGARPARPW BERGGGGGGGARPARPW BERGGGGGGGGARPARPW BERGGGGGGGGGGGGGAGGGGGGGGGGGGGGGGGGGGGGG							
QAPDHFLFAKLKTQTINASDPRIGTSIKYPECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGWYHVITDIQVGSRNYSCTTGHRLIG HSSAECLLSGNTAHWSTKPPICQRIPGLPTII ANGDPISTINENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGRAPQCIIPN KCTPPNVENGILVSDNRSLSLALBVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPFELHGEHTPSHQDNFSRGQEVFYSCEPGY DLRGAASLHCTPQGDWSSEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCEHFCPN PPALLNGRHTGTPSGDIPYGKEISYTCDPHDPR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASSTIPNDFEFPYGTS LNYECRGYGKMFSISCLENLVWSSVEDNC RRKSCGPPPFSNGMYHINDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNYTWDKKAPICEIL SCEPPTISNGDFYSNNRTSFHNGTVVTYQCI GCEPPTISNGFYSNKRTSFHNGTVVTYQCI GSCEPPTISNGFYSNKRTSFHNGTVVTYQCI GSCEPPTISNGFYSNKRTSFHNGTVVTYQCI GSCEPPTISNGFYSNKRTSFHNGTVVTYQCI GSCEPPTISNGCASSHCTVQCQTNGRWGPKLPH CSRVCQPPPELHGEHTLSHQDNFSFGGEVFY SCEPSYDLRGAASHCTVQCTNGRWGPKLPH CSRVCQPPPELHGHTLSHQDNFSFGGEVFY SCEPSYDLRGAASHCTVQCTNGRWGPKLPH CSRVCQPPPELHAGHTLSHQDNFSFGGEVFY SCEPSTDLRGAASHCTVLAGMKALWNSSVPVC EQIFCRNPPALLNGRHTGTPLDDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPYGAACPHPPKIONGHYIGGHYSI VLPGMTISYTCDPGYLLVGKGFICTDQGIWS QLDHYCKEVNCSPPLFMKNGSKELBMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFLTLLILILISWI UKHRKGNNAHENPKEVAHIHISQGGSSVHP RTLQTINENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRGSHAGNLTVAVVLPLANTSY PWSWARNGRAVELALAQVKARPGPSTAHWRYPLL TAGAPALGFGGVKDEYALTRAGFSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRGRLVITYCDPLEAPEDDLSHYT RLATMFRKGRVIYICSSPDAFRILMLALAEA GLCGEDYVFFHLDIFGQSLQGGQGPARRPW ERGDGQDVSARQAFQAAKJITYKDPDNPEYL RLATMFRKGRVIYICSSPDAFRILMLALAEA GLCGEDYYFFHLDIFGQSLQGGGGPARRPW				ł			
EÝYGRPESITCLDNÍ-VWSSPKDVCKRKSCKIP PDPVNGMYNTIDIQVGSNYSCTOFRILIG HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI ANGDPISTINKENFHYGSVVTYRKCNI-GSRGRK VFELVGERSIVCTSNDDLOVGIWSGPAPQCIJEN KCTPPNVENGILVSDNRSLFSLNEVPERCQP GFYMKGPRVKCQALNKWEPELPSCSRVCQ PPEILHGEHTPSHQDNFSPGQEVPYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSL WNNSVPVCEHFCFN PPAILNGRITGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSYRAGHGKKTEQFPFASFIPINDFEPFVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPNGMVHINTIDTGGSTVNYSC NEGGRALIGSPSTTCLVSGNNVTWDKKAFICEI SCEPPTIISNGDFYSNNRTSFHINGTVVTYQCH TGPDGGQLFELVGERSIVCTSDDQVGVWS PPRCISTNKCTAFEVENARVPGNSFFSLTEI IRFRCQPGFVMGSHTVQCQTINGRWGFKLPH CSRVQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSSFPSLTEI IRFRCQPGPMVGSHTVQCOTINGRWGFKLPH CSRVQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSFPARCTV KSCDDFLGQLPHGRVLLPLINLQLGAKVSFVC DGGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCNPPARALNGRHTGTLGDIPYGKEVSYT CDFIPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPYGAACPHPFUNGHYGIGHYSI VLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSPPLFMNGISKELEMKKVYH YGDYYTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTFFLLIILI-SWN ULHRRGGNAHENPKEVAHLHISQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLLLLL LLEPPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAARWSPLIL TAGAPALGFGGVKDFYALTRAGPSYAKLGDF VAALHRRLGWEQALMLYAYRPGDEEHCFF LVGGLFMRVRGRAVITCARPTARHWPULL TAGAPALGFGGVKDFYALTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVGGLFMRVRGRAVITCARPTARHAUNLALEA GLCGEDYVFFHLDIFGGSLQGGGGPARRPW ERGDGQDVSARQAFGAARIITYKDPDNPEYL							
PPPVNGMVHNTIDIQVGSRNYSCTTGHRLIG HSSAECLSGNTAHWSTXPPICQRIPGGLPTI ANGDFISTNRENFHYSOVITYRCNLGSRGRK VFELVGEFSIYCTSNDDQVGIWSGPAPQCIPIP KCTPPNVPNGILUSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPSELLGBHTTSHDDNSSPGOEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHFCPN PPALLNGRHTGTFSGDDPYGKEISTTCDPHEDR GMFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEOPFFASPTIPNDFEFPVGTS LNYECRPGYFGKMFISISCLENLVWSSVEDNC RRKSCGPPPEFFNGMVHINTDTQFGSTVNYSC NEGFRLIGSSTTCLVSGNNVTVDKKAPICEII SCEPPPTISNGDFYSNTSFHINGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENARLVPGNRSFFSLTEI RRRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVQQPPPELHGEHTLSHQDMFSPGGEVFY SCEPSYDLRGAASLACTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPIN.QLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHYLOPHYQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMOISKELEMKKVYH YGDYVTLKCEDGYTLLGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLISTTFILLIIFLSWI LKHRKGRNAHENNEVAHLHSQGGSSVHP RTLQTNEENSRVLP 803 2153 A 6574 2 3233 HGRSARLAVPAEAMPGPRRPAGSRIRILLL LLLPPLLLLRGSHAGNLTVAVVLPLANTSY PWSWARRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVGCVSTAAPLAAVDLKWE HNPAVFLGFGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEVALTTRAGPSVAKLGDF VAALHRRLGWEQALML VAYRRGDEEHCTF LVEGLFMRVRDRINTVOHLEFAEDDLSHYT RLLRTMPRKGRVNICSSPDAFRTIMLLALEA GLCGEDYVFFHLDIFGQSLOGGGPAPRRPW ERGDGGODVSARQAFQAAKIITYKDPDFYYL							l -2
HSSAECILSONTAHWSTKPPICQRIPCGLPPTI ANGOPISTNRENFHYGSVVTYRCNLGSRQRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRVKVCQAINK WEPELPSCSRVCQ PPPEILHGEHTPSHQDDNSSPGOEVTYSCEPGY DLRGASLHCTPQDDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHFCPN PALINGRHTGTPSGDPYGKEISYTCDPHDR GMTFNLIGESTRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPNDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHNITDTOFGSTVNYSC WEGFRLIGSPSTTCLVSGNNVTVDKKAPICEII SCEPPPTISNGDFYSNNRTSFEINGTVVTVQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI RFRCQPGFVMVGSHTVQCOTINGRWGPKLPH CSRVCQPPEILHGEHTLSHQDMFSRGEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGFILGDIPYGKEVSYT CDPHPPRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISTTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSPPLFMOISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTFFILLIELISMI ILKHRKGNNAHENPKEVAIHLHSQGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAVPAEAMFGPRRPAGSRLRILLL LLLEPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWANTOFANELAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRTAHWRYPLL TAGAPALGGGVXDEVALTTRAGFSVAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRINITVUDLEFAEDDLSHYT RLLRTMPRKGRVTICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLOGGGPAPRRFY RLLRTMPRKGRVTICSSPDAFRTLMLLALEA GLCGEDVYFFHLDIFGQSLOGGGPAPRRFY RERGDGOGOVSARQAFQAAKIITYKDPDNPEYL			1	1			l
ANGDFISTNRENHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRIVKCQALNKWEPELPSCSRVCQ PPSELLIGEHTPSHQDDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLPLNLQLGAKVSPVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHFCPN PPALLNGRHTGTPSGDDPYGKEISYTCDPHDR GMFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEFFNGWHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTVQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVYPGRRSFFSLTEII RRFCQPGFVMVGSHTVQCQTHGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPGGDWSPEAPRCTIV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIPCPPPALINGRHTGTPLADIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNOVWS SPAPRCELPVGAACPHPPKQNGHYIGGHVSL YLPGMTISTYCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRW PPLAKCTSRTHDALIVGTLSGTFFILLIIFLSWI LKHRKGNNAHENPKEVAHLHSQGGSSVHP RTLQTNEENSRVLP 803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWANKVGPAVELAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRAFTAHWRVPLL TAGAPALGFGVKDEVYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRGDDEEHCFF LVEGLFMRYRBLNITVVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRILMLLALEA GLCGEDYVFFHLDIFGQSLOGGGPARRFYY RCLAGGGVACHTYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRGDGEHCFF LVEGLFMRYRBLNITVVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRILMLLALEA GLCGEDVYFFHLDIFGQSLOGGGPARRFYY RCLAGGGVACHTYANGLALEA GLCGEDVYFFHLDIFGQSLOGGGPARRFY REGRGGOVAGACAITTYCDPDNPEYL		ļ	Ì				
VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLPSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQGVFYSCEPGY DLRGAASLHCTPGGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGWSSPAPRCE LSVRAGHCKTPEQPFSAPIPENDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEFFNGMVHINTDTGGSTVNYSS NEGFRALIGSSTTICLVSGNNVTVDKKAPICEII SCEPPPTISNGDFYSNNRTSFHINGTVVTVQCH TOPDGEGLPELVGERSIYCTSKDDQVGWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCDTNGRWGPKLPH CSRVCQPPELHGEHTLSHQDNFSFGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPPRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGPIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSFWSQQADDRWD PPLAKCTSRTHDALIVGSTSGTFFILLIFLISMI LK#RKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP 803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWANTVGPAVELAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRTTAHWRYPLL TAGAPALGGWXDEVYALTTRAGPSVAKLGDF VAALHRRLGWEQQALMLYAYPRGDEEHCF LVGGLFMRVRDRINITVVDHLEFAEDDLSHYT RLLRTMPRKGRVIVICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQGSLOGGGPARRFYP ERGDGGODVSARQAFQAAKIITYKDPDNPEYL							1
KCTPPNYENGILVSDNRSLFSLNEVVEFRCOP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPPELLIGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQQDWSFBAPRCAVKSCDDF LGQLPFIGNLVFINLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPALLNGRHTGTPSGDIPYGKEISYTCDPI+DR GMTPTNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPNDFEFPYGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHINIDTOFGSTVNYSC NEGFRLIGSPSTTCLVSGNNYWDKKAPICEII SCEPPPTISNGDPYSNNRTSFHNGTVTVQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENARIVPGNRSFFSLTEI RIFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPALLNGHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGWS SPAPRCELPVGAACPHPFKINGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI LIKHRKGNNAHENFKEVAIHLHSQGGSSVHP RTLQTHESNSTVLY BTRTCTPFELISTLY VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVVKDEYALTTRAGPSYAKLQDF VAALHRELGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRFW		Ì	f	ĺ			
GFWMKGPRRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHGFCPN PALLINGRHTGTPSGDIPYGKEISYTCDPHEDR GMTPNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEGPFFASPTIPINDEFEPYGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPRMGM*HINDTGFGSTVNYSC NGGFRLIGSPSTTCLVSGNNVTWDKKAPICEIL SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAFEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFNNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCADDRWD PPLAKCTSRTHDALIVGTLSGTTFFLLIFLSWI ILKHRKGNNAHENPEPAHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAVPAEANPGPRRPAGSRLRILLL LLLPPLLLLRGSHAGNLTVAVVLPLANTSY PWSWAIRVGFAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW				!			l
PPPELLHGEHTPSHÖDNESPGGEVFYSCEPGY DLRGAASIHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPALLNGRHTGTPSGDPYGKEISYTCDPHEPR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFFASFIPINDEFFFVGTS LNYECKPGYPGKMFSISCLENL.WSSVEDNC RRKSCGPPPEPRIGMVHINTDTOFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDPYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGGRSTYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSFGGEVFY SCEPSYDLRGAASAHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFFLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPALINGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFNINGISKELEMKKYYH YGDVYTLKCEBGYTLLGSPWSQCQADDRWD PLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHKKGNNAHENPKEVAHHLBSQGSSVHP RTLQTINEENSRVLL LLIPPLLLLIRGSHAGNLTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVVDEY ALTTRAGPSYAKLGUP VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFQGSLOGGQGPAPRRFW EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAFQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAKIITYKDPDIPFEYL EEGDGQDVSARQAAKIITYKDPDIPFEYL EEGDGQDVSARQAARAITYKDPDIPFEYL EEGDGQDVSARQAAAKIITYKDPDIPFEYL EEGDGQDVSARQAAAKIITYKDPLEFERDER							l · · · · · · · · · · · · · · ·
DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPAILINGRHTGTPSGDIPYGKEISYTCDPHEDR GMTFNLIGESTRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPINDFEFFVGTS LNYECRFGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGGFYSNNTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFWMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPELLHGEHTLSHQDNFSPGGEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNOVWS SPAPRCELPVGAACPHPPKLQNGHYIGGHVSL YLFGMTISYTCDFGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLLWGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDVVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIFLSWI ILKHRKGNNAHENPKEVAHLHNSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGRRPAGSRLRLLL LLLPPLLLLIRGSHAGGNTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLOPGCVYAAAPVGRFTAHWRYLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMWRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGSQLGGGQGPARRFW ERGDGQDVSARQAFQAAKHITYKDPDNPEYL							
LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPPN PRAILNGRHTGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEGPFASPTIPINDFEFFVGTS LNYECRFGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGGFYSNNRTSFHNGTVYTYQCH TGPDGEQLFELVGERSIVCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVWYGSHTVQCQTNGRWGPKLPH CSRVCQPPELHGEHTLSHQDMSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSPPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAHLHSQGGSSVHP RTLQTNEENSRVLP  ### HGRSARLAAVPAEAMPGPRPAGSRLRLLLL LLLPPLLLLIRGSHAGNLTVAVVLPLANTSY PWSWAIRAVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMU, YAYPROEDEHCFT LVEGLFMRVRDRINTTVDHLEFAEDDLSHFT RLLRTMPKGRVIYICSSPDAFRTLMLLALEA GLCGEDVYFFHLDIFGSQLQGGQGPARRRVW RCGDGQDVSARQAFQAAKIITYKDPDNPEYL			ŀ				
KGSSVSHCVLVGMRSL WNNSVPVCEHIFCPN PPAILNGRHTGTPSGDIPYGKEISYTCDHPPDR GMTFNLIGESTRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEOFPFASPTIPNDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENARVPGNRSFFSLTEI IRPRCQPGFPWGSHTVQCQTNGRWGPKLPH CSRVQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYYLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGSHAGGNLTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGGVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRFGDEEBLCFF LVEGLFMRVORGNITTVHLIFFLAEDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIIFGSPGQGGGAPRRPW ERGDGGDVSARQAFQAAKIITYKDPDNPFYL			1				1 .
PPALLNGRHTGTPSGDIPYGKEISYTCDPHDR GMTFNLIGESTIRCTISDHGNGVWSSPAPRCE LSVRAGHCKTPEOFPFASPTIPINDFEFPYGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSSPTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENARRVPGNRSFFSLTEI RFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGGEVFY SCEPSYDLEAGASHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPALNGRHTGFPLODIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTDPGYLLVGKGFFICTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI LKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTINEENSRVLP RTLQTINEENSRVLP WSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCAPPARAPORFTAHWRVPLL TAGAPALGFGVVAAAPVGRFTAHWRVPLL TAGAPALGFGVVDEVALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYPRGDEEHCFF LVEGLFMRVDRLINITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGGGPAPRRW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			l				•
GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEGPFASPTIPINDFEFFVGTS LNYECRFQFFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEFFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTVQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVWVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSFGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPPKIQNGHYGIGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWAIRVGPAVEJALALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRINTTVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSQGGPAPRRPW ERGDGQDVSARQAFQAAKITTYKDPDNPEYL RLRTMPRKGRVIYICSSPDAFRTLMLLALEA			l				1
LSVRAGHCKTPEQFFASPTIPNDFEFFVGTS LNYECRPGYFGKMFSISCLENLWSSVEDNC RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNODFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIVCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAVPAEAMPGPRRPAGSRLRILLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWAIRVGPAVEJALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVDRRINTVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTIMLLALEA GLCGEDYVFFHLDIFGQSQGGPAPRRPW ERRDDGDOVSARQAFQAAKITTYKDPDNPEYL RERDDGDOVSARQAFQAAKITTYKDPDNPEYL ERRDDGDOVSARQAFQAAKITTYKDPDNPEYL				<b>,</b>			l
LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPRIGMVHINITDTGFGSTVNYSC NEGFILIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGRRSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASHCTVQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILINGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAIENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAVPAEAMPGPRRPAGSRLRLLLL LLPPLLLLRGSHAGNLTVAVVLPLANTSY PWSWA\RVGPAVEALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGFTTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRIGWRQALMLYAYRPGDEHCFF LVEGLFMRVRDRINITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGGPAPRRPW ERGDGQDVSARQAFQAAKITTYKDPDNPEYL				}			l · ·
RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTICLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPELHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSPFLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLIGSHAGNLTVAVVLPLANTSY PWSWALVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKUITYKDPDNPEYL				İ			1
NEGFRLIGSPSTTCLVSGNIVTWDKKAPICEII SCEPPPTISNGDFYSNNTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDQVQVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPPGEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGSHAGNILTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMILYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQFAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			ŀ	ŀ			<b>)</b>
SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVTLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLLGSTWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLL LLLPPLLLLLRGSHAGNLTVAVVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			l				
TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSVLP RTLQTNEENSVLP WSWARVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMILLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			į				
PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFYMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFFLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLK WE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMILALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		ļ					
RFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIVICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			İ		•		TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRINITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLAALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGGQDVSARQAFQAAKIITYKDPDNPEYL			İ				PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLL LLLPPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWARVGPAVELALAQVKARPDLPPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRINTTVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTIMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			]				IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
KSCDDFLGQLPHGRVLLPINLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP RTLQTNEENSRVLP B03 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYPGDEEHCFF LVEGLFMRVDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		1	Ì				CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY
KSCDDFLGQLPHGRVLLPINLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP RTLQTNEENSRVLP B03 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYPGDEEHCFF LVEGLFMRVDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL							SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRPAGSRLRLLLL LLLPPLLLLLRGSHAGNLTVAVVLPLANTSY PWSWA'RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAAAHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			]	]			KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGISHAGNLTVAVVLPLANTSY PWSWAIRVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL				'			DEGFRLKGRSASHCVLAGMKALWNSSVPVC
CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYPPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			ŀ				EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHILHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL							
YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHILHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRGISHAGNLTVAVVLPLANTSY PWSWA'RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYAALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL							
QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHILHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARŁAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYAALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL							
YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		1	1	]			
PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL				1			
ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			ļ	1			
RTLQTNEENSRVLP  803 2153 A 6574 2 3233 HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			t	I			
803 2153 A 6574 2 3233 HGRŠARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			1				
LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL	803	2152		6574	3	2223	
PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL	505	2133	^	05/4		3433	
VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		1	1				
HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		j	1				
TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		1	1	l			
VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			1	1			
LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			1	1			
RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			ŀ	1		,	
GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		1	ŀ	i			
ERGDGQDVSARQAFQAAKIITYKDPDNPEYL		[	ĺ	1			
ERGDGQDVSARQAFQAAKIITYKDPDNPEYL			l				GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
EFLKOLKHI.AYEOFNFTMEDGI VNTIPASEU							
		l	i				ERGDGQDVSARQAFQAAKIITYKDPDNPEYL

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
			'	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	•	/=possible nucleotide deletion, \=possible
ļ				sequence		nucleotide insertion  DGLLLYIQAVTETLAHGGTVTDGENITQRMW
						NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
						NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
						YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV GSLSLLGILIVSFFIYRKMQLEKELASELWRVR
					•	WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL
]						LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
	·				,	IELTRKVLFELKHMRDVQNEHLTRFVGACTD
	!					PPNICILTEYCPRGSLQDILENESITLDWMFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
			:			DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
					!	LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
	,					IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR PSLALOSHLEELGLLMORCWAEDPQERPPFQ
						QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
						EELVEERTQAYLEEKRKAEALLYQILPHSVAE
						QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE STPMQVVTLLNDLYTCFDAVIDNFDVYKVET
				i		IGDAYMVVSGLPVRNGRLHACEVARMALAL
						LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV
						VGLKMPRYCLFGDTVNTASRMESNGEAL\KI HLSS\ETKAVL\EEFGGFELELRGDVEMKGKG
						KVRTYWLLGERGSSTRG
804	2154	Α	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
						MSERVSGLAGSIYREFERLIVRYDEEVVKELIP   LVVAVLENLDSVFAQDQEHQVELELLRDDNE
						QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
						DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT
						KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
						AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
						SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
						QVAQETRNVSTGSAENEEKSEVQAIIESTPEL
						DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
						FEELSSAGSGLIGDVDEGADLLGMGREVENLI   LENTQLLETKNALNIVKNDLIAKVDELTCEK
						DVLQGELEAVKQAKLKLEEKNRELEEELRKA
1						RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
						MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK
						KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP
						GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG
						QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
						VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
1	}	1				TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS
						FTVCNSHVLCIASVPGARETDYPAGEDLSESG
}			[		1	QVDKASLCGSMTSNSSAETDSLLGGITVVGC
1			1			SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS
			]			QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND
						SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
		<u> </u>		· ·	J	MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW
						DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
		Ì	ĺ			YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ
					1	DVDIEPYVSKMLGTGKLGFSFVRITALMVSC
<u> </u>			<del></del>	·		

000 **	CECTO	1 14-4	Loro	Disable 1	Dundiete J J	Amino cold grouper (4-Albeita C. C. )
SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	ļ	ļ	peptide	1	/=possible nucleotide deletion, \=possible
		ŀ		sequence		nucleotide insertion
						NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
	1	1	ł			ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
			ŀ	· ·		FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
	<u> </u>		<u></u>			ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	A	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
		1			İ	SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP
			ľ			VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
			Į.			SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ
	İ	ĺ	ĺ	:		DSGLYACVIRNSTYCMKVSISLTVGENDTGL
		]				CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
						REPEILWYKECRTKTWRPSIVFKRDTLLIREV
			ĺ		}	REDDIGNYTCELKYGGFVVRRTTELTVTAPL
						TDKPPKLLYPMESKLTIQETQLGDSANLTCRA
		1				FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
			Į.			SDI\KILKEHLGEQEVSISLIVDSVEEGDLGNYS CYVENGNGRRHASVLLHKRELMYTVELAGG
			ŀ		1	LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA
	•		1	ļ		EELDGDNKDYDAYLSYTKVDPDQWNQETGE
		1				EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT
	ļ			1		YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF
		ľ	1			ELETRLRNMLVTGEIKVILIECSELRGIMNYQE
				ļ		VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
						LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
	]	ļ		ļ	1	VSAISMAAATSTALATAHPDLRSTFHNTYHS
		1				QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT
		ĺ	]		[	YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
		ļ				ILPLLPRETSISSVIW
806	2156	Α	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL
	Ì					HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT
		j .	1	j	]	QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL
			İ		<u></u>	IQNKYFGDVDIPRAKVVRVCQALMDYKVFE
						AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD
						SQLGKENKLYSPARYADALFKSSDIRSASLED
						LWENLSLKPANSPHVNISTTLSPQVINEVWQE
		1	}	1	l	ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
						RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA
		ŀ				AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE
		I	1	1		
		i	İ		,	LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
					·	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF
					·	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFLMDHQKDVFKIPGT
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFLWDHQKDVFKIPGT LVHKIVSVKVLMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFOTAS\OPOAEFOLT
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFOTAS\OPOAEFOLT
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMILDLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA
807	2157	A	6615	4198	2094	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
]	]			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	ļ <b>'</b>	/=possible nucleotide deletion, \=possible
ľ				sequence		nucleotide insertion
<b> </b>				<del> </del>		LDLSSNPGLEVATGALGGLEASLEVLALQGN
		l	İ			GLMVLQVDLPCFICLKRLNLAENRLSHLPAW
				}		TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR
						RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA
1	ļ	i	Į.	ļ		TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI
		ľ				NLIILTFILVSAILLTTLAACCCVRRQKFNQQ
J		J	j	ļ	}	YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG
						ARHRNNTEKKHPGGGESDASPEAGSGGGGV
				ļ		ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN
	l	i				AGSVGLALIVWIVTGFITVVGALCYAELGVNI
				1	İ	PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP
				· ·		TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA
		1				AICLLLLTWVNCSSVRWATRVQDIFTAGKLL
	ļ	j	]	j	Į	ALALIIMGIVQICKGEYFWLEPKNAFENFQEP
1		I				DIGLVALAFLQGSFAYGGWNFLNY\VTEELV
	ŀ		İ			DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT
1	i	1	1		i ·	AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM
1		i				PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP
i						SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD
		ļ		•	•	MYTLINYVGFINYLFYGVTVAGQIVLRWKKP
	1	ĺ		İ		DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV
1		1				CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI
				1		ELLTLVSQKMCVVVYPEVERGSGTEEANED
			1			MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH
				ĺ		TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY
}		ļ		ļ		LFTSPQRVYRNFHYRKQTKDQWARDDPAFL
	}	}				VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV
		1	l		i	VLIDCVGVGLLIATLMWFISNKYLVKRQSRD
}			ļ	ł		YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN
		ĺ				HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL
			1			GYSVGLLFFS\ALPFLKNTVILLYPFAPLILLYG
	<u> </u>	_	_			LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA
	ĺ	[				QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP
						EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK
		1	Ì		ļ	PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ
ļ	1	}				IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT
		1			ļ	QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF
	1			]	1	IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN
	L	<u>L</u>	<u>L</u>			F
811	2161	A	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG
1		1				AAASISQEPSLADIKARIQDSDEPDLSQNSITG
1		1	ļ	1		EHSQLLDDGHKKARNAYLNNSNYEEGDEYF
1	l·	1	1	İ		DKNLALFEEEMDTRPKVSSLLNRMANYTNLT
[		l				QGAKEHEEAENITEGKKKPTKTPQMGTFMG
		l				VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF
1	1	1	1			AIVLICCCCTMLTAISMSAIATNGVVPAGGSY
1	[		1	[		FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL
	Ì	1		Ī		GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN
				1	}	MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL
1	]	1	}		]	ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT
	1	1			1	LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS
					ļ	QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN
1	1	l	1	}		LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL
	1		1			VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD
1		1	1	ľ	ł	AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV
1	t	I		1		VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS
i	l	l	ł	ł	1	FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV
ſ						

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l noa	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
цепсе	uonce		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciice		ł	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
				,		nucleotide insertion
		<u> </u>	<del> </del>	sequence		FGHSKANGEPTWALLLTAAIAELGILIASLDI.
		į				
						VAPILSMFFLMCYLFVNLACALQTLLRTPNW
		[				RPRFRYYHWALSFMGMSICLALMFISSWYYA
		l				IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS
						LSAARFALLRLEEGPPHTKNWRPQLLVLLKL
						DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV
		Í				GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ
		ļ		ļ		LVVAAKLREGISHLIQSCGLGGMKHNTVVM
		1				GWPNGWRQSEDARAWKTFIGTVRVTTAAHL
						ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG
				1		GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE
		1		]		DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
		ļ	J	J		DISAYTYERTLMMEQRSQMLRHMRLSKTER
		1				DREAQLVKDRNSMLRLTSIGSDEDEETETYQ
		I		İ	İ	EKVHMTWTKDKYMASRGQKAKSMEGFQDL
						LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA
	ļ	1		,		KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL
		İ		1		ERVLLVRGGGSEVITTYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG
				1	• • •	DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP
		1				MEEEEAQPMAAPEGKRSLANGPNAGEQPGO
						VAGADFESEDEGEEFDDWEDDYDYPEEEQLS
						GAGYRVSAALEEADKMFLRTREPALDGGFQ
	1		1	1		MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG
	[			1	ĺ	CDEIIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC
013	2103	^	0030	/08	1333	
	ŀ	İ		1	ļ	WQYRQLSALHRAPRPTRPDKARRLGYKAKQ
			1			GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG
						ATYGKPVHHGVNQLKFARSLQSVAEERAGR
	1			]	]	HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK
	İ	ļ	1		ļ	AIRRNPDTQWITKPVHKHREMRGLTSAGRKS
		l	ļ			RGLGKGHKFHHTIGGSRRAAWRRRNTLQLH
			ļ			RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR
		1	Ì			DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT
	1	i		ļ	1	AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN
	1	l		l	<b>{</b>	PLTKESIRQKEMESKRLRLLQEEDRRKKIARM
	1	Į.		[		GFNASSMLRKSQLGFLNVTNYCHLAHELRLS
	!			I		CMERKKVQIRSMDPSALASDRFNLILADTNS
	[			ĺ	ĺ	DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF
		1		1		MHENLYFTNRKV\NSVCWASLNHLDSHILLC
				1	1	LMGLAETPGCATLLPASLFVNSHPAGIDRPG\
			1		ł	MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR
	1	l	ì	i	l	RVLLTNVVTGHRQSFGTNSDVLAQQFALMA
	-	ļ	]			PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF
		1	1	l		HDSAVTSVRILQDEQYLMASDMAGKIKLWD
	[	1	[	1	[	LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL
	İ	ł		1	1	VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA
				1	1	DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY
			1 .			SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG
		١	55.5	1		PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE
		]	]	}	<b>}</b>	RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT
				1		PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF
		1	l			QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA
			1		1	I COLLIELELAL I FUGLODA VLALATI PKNIKA
			Ì		i	1 '
						LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS
						LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR
						LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV
				}		LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Berice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence		]	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}				peptide	sequence	/=possible nucleotide deletion, \=possible
1	i .			sequence		nucleotide insertion
<del></del>	<del> </del>		<del> </del>	sequence		RVFETLKDLKVLNLAYNKINKIADEAFYGLD
ļ		İ				NLOVLNLSYNLLGELYSSNFYGLPKVAYIDL
		ŀ				QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
1	l					FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR
	ļ					LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ
J	ļ	}	}			TPSENPSLEQLFLGENMLQLAWETELCWDVF
		ŀ	1			EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR
1		[	1			GLSLNSNRLTVLSHNDLPANLEILDISRNQLL
1		l				APNPDVFVSLSVLDITHNKFICECELSTFINWL
i						NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE
1	ĺ	1	ĺ			GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV
						TKFRGFCFICYKTAORLVFKDHPQGTEPDMY
			1			KYDAYLCFSSKDFTWVQNALLKHLDTQYSD
1	1					NYDAYLCFSSKDF1 WYQNALLKHLD1Q1SD   QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
ľ	1	}	l			
1	1	1		Ì		KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL
į		l				NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ
	ĺ					YLRWPEDLQDVGWFLHKLSQQILKKEKEKK KDNNIPLQTVATIS
016	01//		((1)		2011	
816	2166	A	6646	1	3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
1						GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
		l			ļ	GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
		ł				KKGQGTAATGNQATPKTAPATMSTPTILVAT
	ł			ł	ł	AVHAYRYTNGQYVKQGKFGAAVLGNHTTR
						EYRILLYISQQQPVTVARIHVNFELMVRPNNY
						STFYDDQRQNWSIMFESEKAAVEFNKQVCIA
						KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE
					] .	VAYTGWLFQNHVLGQVFDSTANKDKLLRLK
	j					LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA
		ŀ				CAVGSEGVIGWTQATDSILVFEVEVRRVKIA
						KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV
	ļ					VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD
l	1	ł	ł			AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI
		ļ				EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ
				ł		MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA
						VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ
]	1		j	1	]	PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
		1		1		QHNTEIRMAVSKVADKMDHLMTKVEELQKH
					•	SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
			1		1	LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
İ			į	1	1	NLMMEKRNNSLQTATENTQARVLHAEQEKA
1		l	1 .	ł	1	KVTEELAAATAQVSHLQLKMTAHQKKETEL
1					1	QMQLTESLKETDLLRGQLTKVQAKLSELQET
		ļ	}	1	1	SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL
		1		Į.	1	RVEKESLEKNLSERKKKSAQERSQAEEEIDEI
	1	1		1	1	RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS
	1					LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
	1	1		1	1	EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA
					[	QITALTKQNEQHIKELEKNKSQMSGVEAAAS
		1	ļ		1	DPSEKVKKIMNQVFQSLRREFELEESYNGRTI
	1	(	1	i .		LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE
		1			i	EKAEERPRRPSQEQSASASSGQPQAPLNRERP
		1	1	1		ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR
	]		1	}		KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP
		ļ	ļ	1	1	TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN
				1	1	PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE
1		1		,		GDPLALGPESPGEPOPPOLKKDDVTSSTGPHK
	1	1				ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE
		1	1			DELFKGATLKALRPKAQPEEEDEDEVSMKGR
1	1	1	1	Ī	1	PPPTPLFGDDDDDDDDDDDWLG
817	2167	A	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG
				<u>.                                    </u>	1. ***:	

SEQ ID	CEO ID	Mot	CEO	Predicted	Predicted end	Amino soid sequence /A - Alonin - C-Custains
NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noa	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
		1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq- uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	delice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		1	914			
			ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion, \=possible
			<del> </del>	sequence		nucleotide insertion
						KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK
	i	ĺ	1	<b>!</b>		VFHSGTAAKSITKKCEKRSSSWKETELVVVD
						TPGIFDTEVPNAETSKEIIRCILLTSPGPHALLL
						VVPLGRYTEEEHKATEKILKMFGERARSFMIL
	İ	Į	1			IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG
	1	{	1	ĺ		DRYCALNNKATGAEQEAQRAQLLGLIQRVV
			1			RENKEGCYTNRMYQRAEEEIQKQTQAMQEL
	1					HRVELEREKARIREEYEEKIRKLEDKVEQEKR
	}		1			KKQMEKKLAEQEAHYAVRQQRARTEVESKD
			1			GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP
				l .		GISARGLSHEGRKQLAVNLTRVLALYRSILDA
	1	1	1	1	į	YIIEFF\TDNLWDTLPCSWQEALDGLKPPQLA
		1				TMLLGMPGEGEVVRYRSVWPLTLLALKSTA
	1		1			CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR
	ĺ	ĺ	ı			KHVRPKKOHEIRRLGELVKKLSDFT/GLHPGC
			l			RRGLRPG\HLSRFMALGLGLMVKSIEGDORL
	ļ			1		VERAQRLDQELLQALEKEEKRNPQVVQTSPR
		j	1			HSPHHVVRWVDPTALCEELLLPLENPCQGRA
			i	[		RLLLTGLHACG\DLSVALLRHFSCCPEVVALA
		1				SVGCCYMKLSDPGGYPLSQWVAGLPGYELP
		ļ				YRLREGACHALEEYAERLOKAGPGLRTHCY
	1	]	J	1		RAALETVIRRARPELRRPGVQGIPRVHELKIEE
			1	1		YVQRGLQRVGLDPQLPLNLAALQAHLAQEN
		1				RVVAFFSLALLLAPLVETLILLDRLLYLQEQA
		l	ļ			LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG
		1	1			OALSVLETEDS
819	2169	A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA
017	2105	1.	1 0001	00	2000	AAGTAVGDRCERNEFQCQDGKCISYKWVCD
						GSAECQDGSDESQETCLSVTCKSGDFSCGGR
		ł	1			VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC
		ļ	1			SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE
		İ				ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD
		1				ASCEALICGLASEACHSSICILATEA ACDUDED
				1	J	
	1	Ì			į	CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP
				·		CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EMTNLKDIGLYNLRNITRG\airieknadlcyl STVDWSLILDAVSNNYIVGNKPPKECGDLCP GTMEEKPMCEKTTINNEYNYRCWTTNRCQK MCPSTCGKRACTENNECCHPECLGSCSAPDN DTACVACRHYYYAGVCVPACPPNTYRFEGW RCVDRDFCANILSAESSDSEGFVIHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFLK NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPERKRRDVMQVA NTTMSSRSRNTTAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVNMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPPNMESVPLDPS
821	2171	A	6691	106	825	ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC GRVLFRGCGVGHKGQVLMGTFILAQDWLSE
021	21/1	А.		100	<i>- 122</i>	SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK
822	2172	A	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI
						SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH
						VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST
		,				QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF
						TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC
Î				• • • •	1	SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	A	6727	3	4063	PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG
						NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA
						SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS
L	L				L	VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL SDSVMNIFKDRNFDSCCICACNMNIKGADVG LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL FLEDELDIFGKNSDIGQAAERRLM\MCQSTFL PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN FLDYISSNNRQTLPCVSWSYDRVQADNNDY WTECFNALEQGRQYVDNPTGGKVDEALVRS ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQOFH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWFNQPW SGEENDNHSRLKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL
824	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID Not of nucleotide control peptide sequence where the peptide sequenc		L CEC TE	1 3 4 - 4	CCC	Deadista	D	LAmina anid anguaman (AmAlamina Co-Custains
nuclecute coticide sequence   U.S.N   Gostain	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
cotide uence    Sequence		1	hoa				D=Aspartic Acid, E=Glutamic Acid,
sequence    1914   September		1	l				F=Phenylalanine, G=Glycine, H=Histidine,
uence    914   ng no first amino acid residue of peptide pepti	eotide	1 -					
amino acid residue of peptide residue of peptide sequence periode sequence	seq-	uence		1			
residue of peptide sequence   Y=Tyrosine, X=Unknown, *=Stop codon, *possible nucleotide deletion, v-possible nucleotide deleti	uence		1	914			
### peptide sequence    Peptide sequence			1		amino acid	of peptide	
825 2175 A 6735 277 1252 RIMGLEPIGNOMILITYGAFAAFSLMTIAVQ TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRYCCLEORIKGLCQIDHPEDADYE ADTAEYPLAVRASSIPPLSVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASATTAEPSYRXYQRSSSSSSTEPSHSRDA SPVOIRGENTLFSTEISMTISRDPLAATIPT ATYNSDRDNSFLQVHNCIQKENDSLHSNTA NRRTTPV  826 2176 A 6744 3 3177 SDDLRTGLEOVQDAESIALPGVVEVLFYNE EDPDISTADLGDVLQDPCSLEYWDELQKYFV ARFFINLSSKVCELQLPDNLVNDOKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL VTFTALAACTRVDSCTPWFYPSLCVSGPGAH LEFFELCHHLDQLGTAAPQVLQPFVSDRNMPS ELEFMYSSSVCELQLPDNLVNDOKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL AQADCKLLECRNVTMQSVVKPSIFGQMAVS SDVVEKLDCTVIVDSVYNLGGHVVSIDLAT AIQAWQQNKCFEVEELYSHFVICNDTQETL REGQUTDENLLSLSHSHGYSWNSHKSSPQL LHICLEGWCNWWSEPFSVDHAGFTRITIQYR GETASLIKVQQLNGVQGIIGCRGCICSYLSQ SIELKVVQHYIGQDGQAVVRSHFDCLTAKQK KENPELSTERLSTERLSTERLSSLASVILLELLPWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLIKQLATKVHGGTVNQLID EFYOPEKSLQPIWPYNKKDSDRNEQLSQWDS PRKVKLISHNFLYELLFWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALITSKW MCDGNG EVVTLDEBAFVDTERLGAPPGHQKLCOFCIS SMVQQGNQIQIGEDETTINTDTYQTFYRQLSV PRVKLSINFYVERLIFLELPWALITSKSMGDN PRVKLSINFYVERLIFLELPWALITSKSMGDN SSCCWSLISHNKLERIKDDIFKFEVYCKKPESCE SMVQQNGNQIGIDETTINTDTYQTFYROLSV TTEWSDADDNSQGTQVVFLTGFGYVYDV VPQCGTVFTUTVAFGAGPLTATNTAPERIV TTEMSDRNSQGTQVVFLTGFGTVYYDV VPQCGTVFTUTVAFGAGPLTATNTAPERIV TTEMSDRNSLGSRGVLYNSKRHIPALVCQGEKAEPI HHNCPVKMLKENIKDDIFKFEVYCKKPISECE CGDLQLDNQLYNSSHFIFARLVVCQGEKAEPI DTYLPNSRLAGHSTHAGGKQVLVENGTSFVK HSGGTLTSITINATSLARNMRUSLDEEHYM AGALFRAGWVOSLDILGSRASLVRSIGNG VADFFRFELKPARLYVEOTTVVYKICT  BRAQUESTICKT					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
825 2175 A 6735 277 1252 RIMGLEPIGNOMILITYGAFAAFSLMTIAVQ TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRYCCLEORIKGLCQIDHPEDADYE ADTAEYPLAVRASSIPPLSVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASEFYKTRINILSAGFICVSQLASVILLPMGGLCIA ASATTAEPSYRXYQRSSSSSSTEPSHSRDA SPVOIRGENTLFSTEISMTISRDPLAATIPT ATYNSDRDNSFLQVHNCIQKENDSLHSNTA NRRTTPV  826 2176 A 6744 3 3177 SDDLRTGLEOVQDAESIALPGVVEVLFYNE EDPDISTADLGDVLQDPCSLEYWDELQKYFV ARFFINLSSKVCELQLPDNLVNDOKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL VTFTALAACTRVDSCTPWFYPSLCVSGPGAH LEFFELCHHLDQLGTAAPQVLQPFVSDRNMPS ELEFMYSSSVCELQLPDNLVNDOKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL AQADCKLLECRNVTMQSVVKPSIFGQMAVS SDVVEKLDCTVIVDSVYNLGGHVVSIDLAT AIQAWQQNKCFEVEELYSHFVICNDTQETL REGQUTDENLLSLSHSHGYSWNSHKSSPQL LHICLEGWCNWWSEPFSVDHAGFTRITIQYR GETASLIKVQQLNGVQGIIGCRGCICSYLSQ SIELKVVQHYIGQDGQAVVRSHFDCLTAKQK KENPELSTERLSTERLSTERLSSLASVILLELLPWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLIKQLATKVHGGTVNQLID EFYOPEKSLQPIWPYNKKDSDRNEQLSQWDS PRKVKLISHNFLYELLFWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALIKERSLGL SSTQIJEGKGQEKPLQNIEPDLVHHLTQAREE VPSBCAVITSTSLKJCALTKHELPTWALITSKW MCDGNG EVVTLDEBAFVDTERLGAPPGHQKLCOFCIS SMVQQGNQIQIGEDETTINTDTYQTFYRQLSV PRVKLSINFYVERLIFLELPWALITSKSMGDN PRVKLSINFYVERLIFLELPWALITSKSMGDN SSCCWSLISHNKLERIKDDIFKFEVYCKKPESCE SMVQQNGNQIGIDETTINTDTYQTFYROLSV TTEWSDADDNSQGTQVVFLTGFGYVYDV VPQCGTVFTUTVAFGAGPLTATNTAPERIV TTEMSDRNSQGTQVVFLTGFGTVYYDV VPQCGTVFTUTVAFGAGPLTATNTAPERIV TTEMSDRNSLGSRGVLYNSKRHIPALVCQGEKAEPI HHNCPVKMLKENIKDDIFKFEVYCKKPISECE CGDLQLDNQLYNSSHFIFARLVVCQGEKAEPI DTYLPNSRLAGHSTHAGGKQVLVENGTSFVK HSGGTLTSITINATSLARNMRUSLDEEHYM AGALFRAGWVOSLDILGSRASLVRSIGNG VADFFRFELKPARLYVEOTTVVYKICT  BRAQUESTICKT		i	}	1	peptide	•	/=possible nucleotide deletion, \=possible
825 2175 A 6735 277 1252 RIMGLEPRGVOMLTITYGAFAAFSLMTIAVG G TDYWLYSRGVCKTKSVSEMETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHEPEDADYE ADTASYPLAVRASSISPILSVILLIPMGGLCIA ASEFYKTRINILSAGIFFVSAGLSNIGHVYIS ANAGDPSKSDSKKNSVSYSWSFYSFAGLSRIA EMVGVLAVHMFIDRIKQLRATARATDYLQ ASAITEPSYRTYRORISRSSSSSTSTEPSHSRDA SPVGIKGPNTLSTEISMYTLSRDPLKAATTPI ATYNSDENSFLQVHNCIQKENKDSLHSNTA NRRTTPY ATYNSDENSFLQVHNCIQKENKDSLHSNTA NRRTTPY SDDLTFADLDDVLQDPCSLEYWDELQKYFV AFREFILSESKVCELQLEDINLVNDOKKLLVSS DLWRPLVINSSONGADDOSSASESOSSTODEL VYFFALAACTRYDSCETPWFVPSLCVSFQFAH LEFHLCHHDQLGTAAPQVLQPFVSDRMPS ELEYMYSREPHMYLRQWNNGSVCOSIGDFL AGADCKLLERWYTVMGSVWFPSHGCMAVS SDVVEKLLDCTVTVDSVPFNLGQHVVASLNT AGAWGCKCFFVELVFSHVCHOTYGEL AGADCKLLERWYTVMGSVWFPSHGCMAVS SDVVEKLLDCTVTVDSVPFNLGQHVVASLNT AGAWGCKCFFVELVFSHVCHOTYGEL AGADCKLLERWYTVMGSVVFPSHGCMAVS SDVVEKLLDCTVTVDSVPFNLGQHVVASLNT AGAWGCKCFFVELVFSHTVCHOTYGETL RFGQVDTDENILASLHSHQVSWFPSHGCMAVS SDVVEKLLDCTVTVDSVPFNLGQHVVASLNT AGAWGCKCFFVELVFSHTVCHOTYGEL RFGQVDTDENILASLHSHQVSWFPSHGCMAVS SDVVEKLLDCTVTVDSVPFNLGQHVVASLNT AGAWGCKCFFVELVFSHTVCHOTYGEL SKAFETSIVQVFSSNSSITVVCTVLTLEPTS QVQRMVFSSLFSHRVGVARFDCLTAKU LPSTLENRELTBLCVFAKAGDEOWSRDVLG SKAFETSIVQVFSSNSSITVVCTVLTLEPTS QVQRMVFSSLFBRSSLDPDHIHLERSSLGL SSTQIBGKGQEFLQNIEPDLTHRTNGFARFG QVQRWFSSLFBRSSLDPDHIHLERSSLGL SSTQIBGKGQEFLQNIEPDLTHRKSLGL SSTQIBGKGQEFLQNIEPDLTHRKSLGL SSTQIBGKGQEFLQNIEPDLTHRKSLGL SSTQIBGKGQEFLQNIEPDLTHRKSLGLOPHTFFYGEATQHD PRAVKLSINKFYVENTIENTFYGTFYRQLS DWFTFEGERTANTTVGERGFYVALFERGLS UNFIESDERFYVBSAGRITTNTNAFFGETYCKKINESCE SMVQQGIQIGENSTTIENTFYGTFYFQLSV QCSKMGQGLISNKELGHGTYTLTLEEDFSRV BIRNCFVKMLKENIKDIPKFFYYCKKIPSECS SLFCVDLFPSISAFTHRASLLELLTHL DNIFLCVARAGGETPAATFICTCDFSRV BIRNCFVKMLKENIKDIPKFFYYCKKIPSECS SLFCDDFFSFELKFART,VVEDTFYYYTKTL DTYLPNSRLAGHSTHLAGGKQVLIAPKYPGLTV AGABLANDRISSULDEHYM AGALFRAGWVGSLDILGSPASLVRSIGNO VADFFTSELKFART,VVEDTFYYYTKTL DTYLPNSRLAGHSTHLAGGKQVLIAPKYPGTFYAGL SAGLIANNELSLEDEHYM AGALFRAGWVGSLDILGSPASLVRSIGNO VADFFTSELKFART,VVEDTFYYYTKTL SDTYLPNSRLAGHTSTHAGGKQVLIAPKYNGTISFUK HSKGTLTSITIALATSLARMMDRLSLEDEHYM AGALFRAGGVCGLEGGLEGLSLAGLELLALATHL A							
825 2175 A 6735 277 1252 RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCTKSVSENETSKKNEEVMT HSGURTCLEGENKGLCGUIDHPEDADYE ADTAEYPLAVRASSIPPLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLCIA ASEFYKTRINILSAGIFVSAGLSVILLFMGGLVIS ANAGDPSKSDSKNISYSYGWSFYFGALSFILA ENVGVLAVAMPEDREKKQLRATARAKTPTU ASAITEIPSYRYXYQRSRSSSSSTEPHSKDAA SPYGKGPNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLISNTA MRRITIPY RATYNSDRDNSFLQVHNCIQKENKDSLISNTA MRRITIPY BEDPISTADLGDVLQPOSLEVYMDLGKVFV AFREFIALGSVCEQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDYSASSSSOSSCOFDL VYTPTALAACIRVDSCTFWFFYSLCVSFQFAH LEHHLCHHDQLGTAAPQVLQPEVSDRNMPS ELEYMIVSSREPHMYLRQWNNGSVCQEIGPL VYTPTALAACIRVDSCTFWFFYSLCVSFQFAH LEHHLCHHDQLGTAAPQVLQPEVSDRNMPS ELEYMIVSSREPHMYLRQWNNGSVCQEIGPL AQADCKLLECRNYTWGSVVEPSFFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AQAWQONKCFPEUELYFSHFVICADTOTETL REGGVDTDENILASLHSHQVSWRSHKSFQL LHCIEGGWNWW SEPFSVDHAAGTFRITQYR GRTASLIKVQQLNGVQQQQQAVVREHPDCLTAKQK LPSVILENBELTELCVKAKGDEDWSRDVCLE SAFPESVIQVPSSNSSITYWCTVTLTLEPNS QVQQRMIVYSFFIFIRSHPPDIHHLEKRSLGL SEGURGVGQEKQUPDIHDHLASLLNSSKW DQWQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVQQRMIVYSFTFROXSHSYVPVTLTLEPNS QVXCHPSFTFX QVXCKSPECS SHYQQRSSHSYVPVTTLTLEPNS QVXCHPSFTFX QVXCKSPECS SHYQQRSSHSNNELACCTULT QVXCHTATATA QVXCHTATATA QVXCHTATATA QVXCHTATATA QVXCHTATATA QVXCHTATATA QXXCHTATATA QXXCHTATA QXXCHTATA QXXCHTATA QXXCHTATA QXXCHTATA QXXCHTATA QXXCHTATA QXXCHTATA QXX			<del></del> -		Boquence		
### 2175 A 6735 277 1252 RMGLFPGROVØHLTTVGAFAAFSLMTIAVG TDYWLYSROVCHTSVESEMTSKNEEVMT HSGLWRTCCLEGNFRGLCKQIGHFPEDADYE ADTASYFLAVARSSIPPLSVILLFMGGLCA ASEFYKTRHJILSAGIFFVSAGLSNIIGHVYIS ANAGDPSKSDSKNISYSYGWSFYFGALSPILA EMVÖYLVAHMFIDRIKRQIRATARATIDTLQ ASAITKIPSYRYRYQRSRSSSSRSTEFSHSRDA SPVGIKGFWILFSTEISMTLSROPLKAATIFT ATYNSDEDNSFLQVHNCIQKERKDSLHSNITA NRRTTPV  ##################################					}		1 7
### TOPYWLYSROVCKTKSVSENETSKKNEEVM HSGU WRTCCLEGFREGLCKQIDHFEDADYE ADTAEYFLRAVRASSIPPLSVILLFMGGLOVIA ASEPYKTRHNILLSAGIFFYSAGLSNIGIUVIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFILL EMYGVLAVHMFDRHKQLBATARATTDYLQ ASAITRIPSTRYTRYTRYGRSSSSSTEPSHSRDA SPOGKGPFILTESTEISMYTLSRDYLAATIPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPY  ##################################		0.55	<del> </del>	(50.5		1050	<u> </u>
HSGL.WRITCLEGNFKGLICKQIDHFPEDADY ADTAEVFLAVARSSIPPLISVILLFMGGLICIA ASEPYKTRINIILSAGIFFVSAGLSNIIGIUVIL ASEPYKTRINIILSAGIFFVSAGLSNIIGIUVILA ANAGOPSKSDSKKNYSVSWGSYSYFGALSFIIA EMYGVLAVEMFEDEHKQLRATARAITDYJJ ASAITREYNYRYRYGRSKSSSSSFYFGALSFIIA BAYGUKGFNTLPSTEISMYTLSRDPLKAAITTY ATYNSDRDNSFLQVINCOKENKDSLHSNITA NRRTTPY  826 2176 A 6744 3 5177 SDDLRTGTFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLITLVRTPYPFNIT EDPDISTADLGDVLQDFCSLEVWDELQKLYVS AREFNLSSSKYCELQLPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTOKTOK APPENLSSSKYCELQLPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTOKTOK APPENLSSSKYCELQLPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTOK LEFHILCHILDQLGTAAPQYLQPFVSDRNMYS ELEYMIVSRFEPHWYLRQWNNGSVCQEIQFL AQADCKLLECRNYTMQSVVKPFSIFGQMAVS SDVVKKLLDCTVIDSVFVNLGQRGGICSVISQ SVVKKLLDCTVIDSVFVNLGQRGGICSVISQ SVVKKLLDCTVIDSVFVNLGQRGGICSVISQ SVVKKLLDCTVIDSVFVNLGQRGICSVISQ GRTASLIKVQQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGEDGGICSVISQ SIELKVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGEDGGICSVISQ SIELKVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGEDGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGEDGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGENGGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGENGGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGENGGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGENGGICSVISQ SIELKVVQHYIGQDQAVVREHEDCLTAKQK LSYMILENETLELCVVALGGENGGICSVISQ SIELKVVQHYIGQDQAVREHEDCLTAKQK VADFFRINGROPGAVGVQCICCGCICS SMYQQQIQIQECKTTINNTPYQIFYKRCDSDRAV DLWLFGGKKVQLVQAGKIIPPQAFQGIG WANTNTVIKSVALKLVHNLTSFRWKDGGIGG EVVTTLDEARVDTERILGAFPOHQKLCQFCIS SMYQQQIQIQECKTTINNTPYQIFYKPGCAC SMYQQQIQIQECKTTINNTPYQIFYKPGCAC SMYQQQIQIQECKTTINNTPYQIFYKPGCAC HENGROPGATAGATATATATATATATATATATATATATATATATAT	825	2175	Α	6735	277	1252	
ADTAEYFLRAVRASSIPPUSVILLFMGGLOVIS ASAGDPSKSDSKKNSYSYGWSFYFGALSHIGH ASEPYKTRINHULSAGDIFYSAGLSNIGHVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSHIGH EMYGVLAVIBMEDHIKQLRATARATIDYIA ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA SPVGIKGTVILSRDPLKAAITITY ATYNSDRNSFLQVINKCIQEENKOBLISNITA ANRATTPY  SDDLRTGIFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLILVRITPYPPTNIT EDPDISTADLGDVLQPCSLEEWDBLQKVFV AFREFNLSSKVCELQLPDDNLVNDQKKLVS UNWILVINSSQNGADDQSSASSESGSSTCDPL VYPTALAACTRVDSCFTPWFVSLCVSFGRAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSTREPHMYLRQWNNGSVCQSICDPL AQADCALLECRNYMGSVVKPSIFGQMAVS SDVYBKLLDCTVINDSVPXNLGQHVVHSLNT AQAWQNNCFPVELVSHFFVCNDTOFETL REGQVJTDENILLASLHSHQYSWRSHKSPQL LHICEBGWGNWRWSEPFSVDHAGTFIRTOYS GRTASALIKVQQUNCYCQMCGRQGICSYLSQ SIELKVVQHYIQQDQQAVVEHFDCLTAKQK ESKAPFYSIVQVPSSNSSIIYWCTVLTLEPSN QVQQRMIVFSPLFMSSHLPPDIHLEKRSLGL SETQIPPGKGQEKPLQNIEPDLVHHLTFQARE YDPSDCAPPITSILKQJATKVTHLEELPWHGGTVNQILD EFYGFEKSLQPIWPYNKKDBRNEQLSQWDS PMKVKLSIWKPYVRTLLELLPWHGGTVNQILD EFYGFEKSLQPIWPYNKKDBRNEQLSQWDS PMKVKLSIWKPYVRTLLELLPWKGGGNG EVYTLDEGAFVDTEIRLGAPPGHQKLCQCIS SMVQGGIQUEDKTTINNTPYQIFYRQLSV DLWLFEGBKVLQVPAGRIPPHTQEAPGIGITY WANTTYHRSVARLVHNLTSFKWGGGNG EVYTLDEGAFVDTEIRLGAPPGHQKLCQCIS SMVQGGIQUEDKTTINNTPYQIFYRPQLSV CNPHSGKEYTRYPDSATFSICPGGQPAMKSS SLPCWDLMPDISQSVLDASLLQKQMLGPSPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLLTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYYLKTLSEDPSRPA PGADSSQCWSLFAHVREPFRQSVAVPLGRFR ENGFCTRAVILTYGEHLGVYLVILAGABAILKUT TIMMFTQSSYPDCKTRAVILPSLLUKVSBGGNG HHELTYHQUSSYPDCKTRAVILPSLUKVSGGNG CSULDDRESFERLERALLYVEDTYVYYKTL ARADTPLYSTNOFREGERLGGLTGLGSASLVRSIGNO VALDFFFERLEVALTFELVYVYKTL ARADTPLTSTVEREGERLGGLTGLGSLGSLGSLA		i	l	l	İ		
### ASEPYKTRINIII.SAGIFFYSAGL.SNIGIGIVID.  ### ANAGEPSKSDSKKNSYSYGWSYFGALSFINA ANAGEPSKSDSKKNSYSYGWSYFGALSFINA ANAGEPSKSDSKKNSYSYGWSYFGALSFINA ASATTRIPSYRYRYQRSRSSSRSTEPSHSDA BYVGIKGPTIPSTEISMYTLSRDPLKAATTPY ATYNSDRDNSFLQVINCIQKENKDSLHSNTA NRTTPY  ### ATYNSDRDNSFLQVINCIQKENKDSLHSNTA NRTTPY  ### SDDLRTGLFQDVQDAESLKLPGVYYEVLFYNN TEDCPGMML WRYPEPRGLTLVHIPVPFNTT EDPDISTADLGDVLQDPCSLEVWDELQKVY ARREFNLSESKVCBLQLPDDILVNDQKKLVS DLWRIVLNSSQNGADDQSSASESGSGSTGDPL VTPTALAACTRUSDCFTPWFVPSLVSPQPAH LEFHLCHILDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCVESQPAH LEFHLCHILDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCVESQPAH LEFHLCHILDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCVESQPAH LEFHLCHILDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYSHQSVVKPSIFGQMAVS SDVVSKLLDCTVIVDSVYFNLGQHVVHSUNT ANQQNKGPEELVSHFVTCNDTQETL RRGQVJTDENILASLHSHQYSWRSHKSPQL LHICCEGWGNWRWSEPFSVDHAGFTFRTQYR GRTASLIKVQQLNGVQKQHIGGRQIICSYNG SIELKVVQHYIGQDGQAVVREHPCLTAKQR LFSYILENNELTELCVKAKQDEDWSRDVCLE SKAPPSVINQVPSSNSIITYWCTLLEPNS QVQQRMIVFSPLFIMSHH.PDPIIHH EKRSLGL SETQUPGKGGEKPLQNHPDLYHHLTPGARE SKAPPSVINQVPSSNSIITYWCTLLEPNS QVQQRMIVFSPLFIMSHH.PDPIIHH EKRSLGL SETQUPGKGGEKPLQNHPDLYHHLTPGARE SKAPPSVINQVPSNSSIITYWCTLLEPNS QVQQRMIVFSPLFIMSHH.PDPIIHH EKRSLGL SETQUPGKGGEKPLQNHPDLYHLTEPNS QVQQRMIVFSPLFIMSHH.PDPIIHH EKRSLGL SETQUPGKGGEKVTLQVPAGKIIPPNTQEAFQIG WANNTNYHKSVAIKJHNILTSPRWCDGGNG EVYTLDEBAFVDTEIRLGAFPGHQKLCQFGLS SMYQQGIQIQUEDKTTIINNTPYGEAFQIG WANNTNYHKSVAIKJHNILTSPRWCDGGNG EVYTLDEBAFVDTEIRLGAFPGHQKLCQFGLS SMYQQGIQIQUEDKTTIINNTPYGFYNKYLG SMYQQGIQIQUEDKTTIINNTPYGFYNVLOG UTTEWBALDIDSSYDCCTKULPSILLRVEPLGE TREMPTIQLSLAYDPYLRKRAELLRLTL DNIFLCVAPGAGPLTNTNRAPEKEL TREMPTIQLSLAYDPYLRKRAELLRLTL DNIFLCVAPGAGPLTNTNRAPEKEL CGGLQLDNQLTNKSNFHFAVLVQGEKAEPI QCSKMGSLLISNKELEEYKECFIKLGTLASAL KILL TRIMPTIQLSAYDVLLVSHASAELLRLTL DNIFLCVAPGAGPLTSPRAGLVITALAAHIY AAGALFRAGWVQSLDILGSPASLVRSING KSILCDNIFLSFSTYRRGPITTARQLVIALAAHIY AAGALFRAGWVQSLDILGSPASLVRSING VADFFTLYYGKTUR AAGALFRAGWVQSLDILGSPASLVRSING VADFFTLYYGKTUR HISKGTLTSITNLATSLARNMDRLSLDEHNYN RQEEWRQLPESSLGGGLAGGLISSLLGAL				1	ļ		HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE
ANAGDPSKDSKKNISYSYGWSFYFGALSFINA EMYÖVLAVIHÖRIRKIQLARATAITDYLQ ASAITRIPSYRYRYQRASSSSSSTEPSHSRDA SPYGIKGPTILPSTEISMYTLSPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRITTPY  BEDERTGLFQDVQDAESLKLPGVYEVLFYNB TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNTWMPS ELEYMIVSREPHMYLRQWNNGSVCQIOFL AQADCKLLEGRIVTNGSVVFPSUGOHSPL TYPTALAACTRVDSCFTPWFVPSLCVSFGAH LEFHLCHHLDQLGTAAPQVLCYPSUSTMMPS ELEYMIVSREPHMYLRQWNNGSVCQIOFL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLEGRIVTNGSVFPSUDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIGCRQIICSYLSQ SDVVEKLLDCTVIVDSVFVNLGGHVVHSINT AQQQGMMVTSPLTLCVKAKQEDEDWSRDVCLE SKAPEYSTVQVPSSNSSITVVWCTVLTLEPNS QVQQGMMVTSPLTMSSHLPDPIIHLEKRSLGL SETQIIPGKQGEKPLQNIEPDLVHHLITQAREE YDPSDCAVPISTLIKQAIKVHPGGTVNQLUD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLIELILFARS DLWFTLEPNS QVVTLDEAFDDTERLGAFPGHQKLCOFCIS SMVQQGIQIDGDEXTTINNTPAWALINGSKW DLWLFEGEKIVLQVPAGKIIPPNIQAEPA PGADSSQCWSLPAVPRFFREYSVAVPLGFGF ENGCTRAIVLTYQEHLGYTYLTLSEPDSRV HHRLYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPTYLTLGFGYVYVDV VHQCGTVFTTVAPEGKAGPILTNTTNRAPEKIV TFKMPTIQLSLAVPDDLTHRKEPVYCKKLIPSCCS HHELYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPTYLYTLEFDFSRV HNRCPVKMLIKENINDIPKEYVYCKLIPSCCS HHELYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPLYQCGGEAPH QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSLILLAVPOKLTRAKASALLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFITHASAGLLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFITHAYCYCGGEAEPH QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSLILLAVPRILARRADILSLJGEHYY RQEGETRRQLPFSLLOFGLRGGSRGNOLSLGEHYN RQEGETRRQLPFSLLOFGLRGGGSRGNIGSHSVK HISKGTLTSTITHALTSLARMDRISJDEEHYN RQEGWRRQLPFSLGGGRQGLSKGIGISLLGAU VADPFRLEYFEGLGGGGRQGLSKGIGISLLGAU RGCGETRRCDFSLGGGGGGAGLSSIGGNO							ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA
ANAGDPSKDSKKNISYSYGWSFYFGALSFINA EMYÖVLAVIHÖRIRKIQLARATAITDYLQ ASAITRIPSYRYRYQRASSSSSSTEPSHSRDA SPYGIKGPTILPSTEISMYTLSPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRITTPY  BEDERTGLFQDVQDAESLKLPGVYEVLFYNB TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNT TEDCPGMMLWRYPERFGLTLVRITEYPSNTWMPS ELEYMIVSREPHMYLRQWNNGSVCQIOFL AQADCKLLEGRIVTNGSVVFPSUGOHSPL TYPTALAACTRVDSCFTPWFVPSLCVSFGAH LEFHLCHHLDQLGTAAPQVLCYPSUSTMMPS ELEYMIVSREPHMYLRQWNNGSVCQIOFL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLLEGRIVTNGSVVFPSUGOHSPL AQADCKLEGRIVTNGSVFPSUDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIGCRQIICSYLSQ SDVVEKLLDCTVIVDSVFVNLGGHVVHSINT AQQQGMMVTSPLTLCVKAKQEDEDWSRDVCLE SKAPEYSTVQVPSSNSSITVVWCTVLTLEPNS QVQQGMMVTSPLTMSSHLPDPIIHLEKRSLGL SETQIIPGKQGEKPLQNIEPDLVHHLITQAREE YDPSDCAVPISTLIKQAIKVHPGGTVNQLUD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLIELILFARS DLWFTLEPNS QVVTLDEAFDDTERLGAFPGHQKLCOFCIS SMVQQGIQIDGDEXTTINNTPAWALINGSKW DLWLFEGEKIVLQVPAGKIIPPNIQAEPA PGADSSQCWSLPAVPRFFREYSVAVPLGFGF ENGCTRAIVLTYQEHLGYTYLTLSEPDSRV HHRLYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPTYLTLGFGYVYVDV VHQCGTVFTTVAPEGKAGPILTNTTNRAPEKIV TFKMPTIQLSLAVPDDLTHRKEPVYCKKLIPSCCS HHELYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPTYLYTLEFDFSRV HNRCPVKMLIKENINDIPKEYVYCKLIPSCCS HHELYHQISSYPDCKTKDLLPSLLRVPLDE YTTEWSDADIDNSQGTQVVPLYQCGGEAPH QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSLILLAVPOKLTRAKASALLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFITHASAGLLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFITHAYCYCGGEAEPH QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSLILLAVPRILARRADILSLJGEHYY RQEGETRRQLPFSLLOFGLRGGSRGNOLSLGEHYN RQEGETRRQLPFSLLOFGLRGGGSRGNIGSHSVK HISKGTLTSTITHALTSLARMDRISJDEEHYN RQEGWRRQLPFSLGGGRQGLSKGIGISLLGAU VADPFRLEYFEGLGGGGRQGLSKGIGISLLGAU RGCGETRRCDFSLGGGGGGAGLSSIGGNO			ļ				ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS
EMYGVILAVIHMEDRIKQIRATARAYIDYLQ ASATRIPSYRYRQRSRSSESTESHSINDA SPYGIKGPNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV SDDLRTGIFQDVQDAESIKLPGVYEVLFYNE TEDCOGMM.WRYPERGILT.VRITEVPPNTT EDPDISTADLGDVLQDPCSLEYWDELQKFV AFREFINLESSKVELQLPDNU-NDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTOPL VTTALAACTROSCFTPWFYSLCVSFQPAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSREPHIMYLRQWNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFQUMA VS SDVVEKLLDCTIVIDSFVNLGQHVVHSLNT AIQAWQQNKCPFVSELVFSIFVICNDTQETL RFGQVDTDENLLASILHSYSWRSHKSSPQL LHICEGWGNWW SEPSPSVHAGTFIERTQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKLVQHYYIGQDGQAVVREHDCLTAKQK LPSYLENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVQVPSSNSSITVVATLTEPNS QVQQRMIVFSIFIRMSH.PDPIIHLERSLIGL SETQIPIGKGGGEPLQNIEPDLHILTGVAREE YDPSDCAVPISTELIKQLATKVEPGGTVAQILD EFYGPEKSLQPIWPYNKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLIELILPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNQEAFGIGIY WANTITYHKSVAIKLVHALITSPKWKDGGNG EVVTLDEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIGEXTTIINLTSFEWKDGGNG EVVTLDEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIEDEXTTIINLTSFEWKDGGNG EVVTLDEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIEDEXTTIINLTSTSEWSDGMMS SIEVENDLMPDISQSVLDASLLVTLSEPPSRPV CNPHSGKEYFRVPDSATTSICPGGEQPAMKSS SIEVENDLMPDISQSVLDASLLVTLSEPPSRPV HENCEYKMLIKENIKDIPKFEVYCKKIPSECS HHELYTQBILGTYTLTSEPPSRPV HENCEYKMLIKENIKDIPKFEVYCKKIPSECS HHELYTQBILGTYTLTSEPPSRPV HENCEYKMLIKENIKDIPKFEVYCKKIPSECS HHELYTQBILGTYTLTSEPPSRPV UPQCGTVFITVAPEGKAGPLTNTINKARAPERIV TFKMFTIQLSLAVPDDLTHRASAELJRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDQLYNKSNFHEAFLYCYGGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCONNEFSERLEPARLYCCTGGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCONNEFSERLEPARLYCCTGGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCONNEFSERLEPARLYCCTGGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCONNEFSERLEPARLYCCTGGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCONNEFSERLEPARLYCCTGGERAEPI CGSKMCGLIFTSTITALATSLARMDRISLSLDEHYN RQEEWRRQLPEGLIFGGGRAGLSSLGSIGMO VADPFRLIPTEGLIFGGGRAGLSSLGSIGMO VADPFRLIPTEGLIFGGGRAGLSSLGSIGMO VAPFRLIPTEGLIFGGGRAGLSSLGSIGMO VAPFRLIPTEGLIFGGGRAGLSSLGSIGMO VAPFRLIPTEGLIFGGGRAGU			i		1		
ASAITRIPSTRIMYLORRSRSSRSTEPSHRDA SPYGIKGFNILPSTEISMYLSPELKATTPY ATYNSDRNSFLQVHNCIQKENKDSLHSNTA NRRTTPV  SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCOGMM.WRYPERGLTLVRITPYPNTT EDPDISTADLGDVLQPPCSLEYWDELQKVFV AFREFNILSESKVCELQLPDNILNYDQKKLVSS DLWRIVLNSSQNGADDQSSASEGSGSTCDPL VTFLALACTRVDSCFTFWFVPSLCVSFQAH LEFHLCHHLDQLGTAAPQVLQPVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIGFL AQADCKLLECRNVTMQSVVKPFSIFQGMAVS SDVVEKLLDCTVIVDSVFYNLQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQEIL RFGQVTDTEBLLASLHSHGVSYWSSHSSPQ. LHICEGWGNWRWSEPSVDHAGTFIRTQYR GRTASLIKVQQHNGVQKQIICGKQIICSYLSQ SIELKVVQHYIGQDGQAVVREHPDCLTAKQK LPSYILENNELTELCVKAKGDWRSDVCLE SKAPEYSIVQVPSSNSSITYVWCTVLTLEPNS QVQQRMVFSPLFINRSHLPDPIIHLEKRSLGL SETQIFIGKGQEKPLQNIEPDLVHHLTPQAREE YDPSDCAVPISTSLIKQLATKVHPGGTVAQILD EFYGPEKSLQPWFYNKKDSDERQLSQWDIS PMRVKLSIWKPYVRTLIELLPWALLINESKW DLWLFEGEKIVQVPAGKIIPPNFQEAPGIGTY WANTNTVHKSVARLVHNLTSPKWKDGGNG EVVTLDEAFAVDTERLGAFPOKLCOFGLS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLJAKQKQMLGFSPA PGADSSQCWSLPAVRFEFFRQSVAVFLGNFR ENGFCTRAIVLTYQEHLGYTVTLTSEPSSFRV IHNRCPVKMLIKENKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTROLLFUTDLEV TFRKMPTIQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHHAAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCNDNFFSEFKLKPARLYVETYVYTKLT- DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVPPVKLRKLVIOPPVALISHASLKLJ AGADLFRAGWVSTARSHARDRUSLDEEHYN RQEEWRRQLPSELGEGLINGLISHASLKJ AGAALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISLACHAH AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAISHASLKJVALAAHY AAGALFRAGWVGGLSHAIGSLGAH VAPPFELFYEGLTRGPGAFVSVSSRGTTSFVK HISKGTLTSITHLATSLARIMADRISLDEEHYN RQEEWRRQLPSELGEGERGAGLSRLGISLLGAISLAGERL			]	j	J		
826 2176 A 6744 3 5177 SDDLRTGLQNENCIQKENKDSLHSNTA ATNSRDNSRJQVHNCIQKENKDSLHSNTA NRRTTPV  SDDLRTGLQDVQDAESLKLPGVYEVLFYNE TEDCRGMM.WRYPERRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKFV AFREFINLESSKVCEQLQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL VTPTALAACTROSCFTPWFYSLCVSFQFAH LEFHLCHILDQLGTAAPQYLQPFVSBTMMPS ELEYMVSFREPHMYLRQWNNGSVCQEIGFL AQADCKLLECRNVTNQSVVKPPSIFGQMAVS SDVVEKLLDCTVIVDSFVYNLGGHVVHSLNT AIQAWQQNKCPEVELVFSHFVICNDTQFTL RFGQVTDENILLASLISHQYSWRSHKSPQL LHICEGWGRNW WSEPFSVDHAGTFIRITQYR GRTASLIIKVQQLNGVQKQHICGRQHICSYLSQ SIELKLVVQHYIGQDQAVVRHDCLTAKQK LPSYLENRELTELCVKAKGDEDWSRDVCLE SKAPEYSTIVQYPSSNSSITVWCTVLTLEPNS QVQQRMIVFSFLFIMRSHLPDPIHLHERSLGL SETQIPGKGQGEPLQNIEPDLYHHLTGAREE YDPSDCAVPISTSLIKQGATKVPPGGTVAQILD EFYGPRESLGPHYPNKKDSDROLSOWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFGGEKVILQVPAGKHIPPNTQAFAGHT WANTNTHKSVARLVHVLTSFKWKDGGNG EVVTLDEAFYDTEIRLGAFFGHKQLCQFGIS SMVQQGQIQENKTTINNTPYQIFYXPQLSV CNPHSGKEYFRVPDSATTSICFGGEQPAMKSS SLFCWDLMPDISQSVLDASLLGKMLGFFFRQSVAVFLGFR PGADSSQCWSLFAIVRFEFFRQSVAVFLGFR ENGFCTRANLYTQBHLGWYTLTSEPSSRV HPHRCZVKMLIKENIKDIPKFEVYCKKLFSECS HHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADDINSQGTQVVPTLTGFGYVYVDV VHQCGTVFITVAFEGKAGPLTNTINRAPEKIV TFKMFITQLSLAVPDDLTHRKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVBIC CGDLQLDNQLYNKSNFHERAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKFKLCTILNEG KSICIONEFSESLEPARL,VEVSHASLKLY ASDHTPLSFSVFERGPFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVKSIGMO VADFFRLPYEGLTRGPGAFVSVSVSRGTISFVK HISKGTLTSITINLATSLARRMDRISLDEEHYN RQEEWRRQLPEGLIGERGCAFVSVSSKGTTSFVK HISKGTLTSITINLATSLARRMADRISLDEEHYN RQEEWRRQLPEGLIGERGCAFVSVSSKGTTSFVK HISKGTLTSITINLATSLARRMADRISLDEEHYN RQEEWRRQLPEGLIGERGCAFVSVSSKGTTSFVK HISKGTLTSITINLATSLARRMADRISLDEEHYN RQEEWRRQLPEGLIGELGGLAGLGISLLGAH VADFFRLEFEGLIGERGCAFVSVSVSRGTTSFVK HISKGTLTSITINLATSLARRMADRISLDEEHYN RQEEWRRQLPEGLIGEGLGGLAGLGISLLGAH VADFFRLEFEGLIGEGLGGLAGLGISLLGAH RGEGEWRRQLPEGLIGGLGAGLGGLGISLLGAH RGEGEWRRQLPEGLIGGERGAE		1		1	ł		
826 2176 A 6744 3 5177 SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCFQMML WYPERGLTLVRITYPFNTT EDPDISTADLGDVLQDFCSLEYWDELQKVFV AFREFNLSESKVCELQLPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL VTTALAACTRVDSCFTFWFYPSLCVSFQFAH LEFHLCHHLDQLGTAAPQVLQFFVSDRIMPS ELEYMTYSREPHMYLRQWNNGSVCQEIQFL AQADCKLLECTVITMGSVVKFSIFGQMAVS SDVVEKLLDCTVITMSSPVDLAKGOFQVVCLE SKAPEYSIVQOPSOSNSSITVWCTVLTLEPNS QVQQRMIVFSFLFIRMSHLPDVHIHLEKRSIGL SFTQIFGKQGERPLQNIEFDLVHHLERSIGL EFYGPEKSLQPIWFYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVTTLLELLPWALLINESKW DLWLFFGGGKVLQVPAGKTSDRNEQLSQWDS PMRVKLSIWKPYVTTLLELLPWALLINESKW DLWLFFGGSKVLQVPAGKTSDRNEQLSQWDS SMVQQGIQIQEDKTTINNTPYQIFYKPQLSV CNPHSKCEYFRVPDSATFSICQFAMKSS SLCWDLEMPDISGSVLDASLQGCPAMKSS SLCWDLEMPDISGSVLDASLQGCPAMKSS SLCWDLEMPDISGSVLDASLQGCPAMKSS SLCWDLEMPSISSVPCKKTKDLIPSLLLRVEPLDE VTTEWSDADDINSQGTQVVPLTGFGVYVVD VHQCGTVFTVAPEGKAGPLTNTNRAPEKIV TFKMPTTQLSLAVFDDLTHHKASAELLRIL DNIFLCVAFGAGFLFGEEPVAALTELTVCVEIC CGDLQLDNQLVNKSNHFHAVLVCQGEKAEPI QCSKMQSLLISNKELEEVKEKCFTKLTINAG KSLCDINEFSELFRARINGLEFTYNSTLTVKTLTLDTVTLTVKTLTLTVCVEIC CGDLQLDNQLVNKSNHFHAVLVCQGEKAEPI QCSKMQSLLISNKELEEVKEKCFTKLTINAG KSLCDINEFSELFRARINGLSCHYYN AGALFPAFSVFERGPFTTARQLVHALAMHY AAGALFRAGWVVGSLDLIGSPASLVRSIGNG VADFFRLPYEGLTRCFGAFVSCVSRGTTSFVK HISKGTLTSTINLATSLARNDMALSLLDEHYN RQCEFWRRQUPESLGCELGRGLSRLGISLIGAH VAGFERFPYEGLTFSTGGLGGLSRLGISLIGA	ĺ			Ì	1		1
REZÓ 2176 A 6744 3 5177 SDDLRTGLFÓDVÓDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEY WEDLGKVEV AFREFINLESSKVEGLQPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSTCDPL VTTALAAGATRUDSCTPTWFVPSLCVSFOFAH LEFHLCHHLDQLGTAAPQVLQPFVSDRNMPS ELEYMTVSREPHMYLR QNWNNGSVCQEIOFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVELLVFHVICNDTQETL RFGQVDTDENILLASLHSHQYSWSHKSSPQL LHRICEGWGNWRWSPSPSDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQVIFGDQGQAVKEHPDCLTAKQL LHRICEGWGNWRWSPSPSSDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQVFIGODQGAVKEHPDCLTAKQL LHRICEGWGNWRWSPSPSSDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQVFISTOLKQVALIFICDLTAKQL SKAPEYSIVQVFSSNSSIIYVWCTVLTLEPNS QVQQRMIVTSPLFIMRSHLDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHAILTPQAREE YDPSDCAVPISTSLKQVATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDNRDCLSQWDS PMRYKLSIWFYYNTLIELLPWALLNTSSKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAPQIGIY WANTNTVIKSVAIKLVHNITTSFK WEDGGNG EVVTLDEEAFVDTERIT,GAPFGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTFYQIFSPA PGADSSQCWSLPAINVREPGSVAVPLGFGFA PGADSSQCWSLPAINVREPGSVAVPLGFFA ROGDSSVLDASLLQKQMLGFSPA PGADSSQCWSLPAINVREPGSVAVPLGFFA PGADSSQCWSLPAINVREPGSVAVPLGFFA ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNKCPVKMLIKENIKDIPKFEVYXCKIPSECS HHELYHQISSYPDCKTKDLIPSLLLRVEPLDE VTTEWSDAIDINSQCTQVVFLTGFGYYYDV VHQCGTVTIVAPGKAGVALIFELXCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEVKEKCFIKLTINEG KSLCDNEFSFELKPARLYVEDTTNYYKIKLT DTYLPNSRLAGHSTHLSGGKQVLYELCTINEG KSLCDNEFSFELKPARLYVEDTTNYYKIKLT DTYLPNSRLAGHSTHLSGGKQVLYBLALAMHY AAGALFRAGWVVGSLDILGSGASLIPSINKLUEHYNALAMHY AAGALFRAGWVVGSLDILGSGASLVSHALAMHY AAGALFRAGWVVGSLDILGSGASLLIPSINK JEGEWRRQDVFSLLGGLGRASLLIPTIV ASDHTPLSTSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSGASLLIPSINK AGALVPRYKLRLVLIQPVILLVSHASLKLYI ASDHTPLSTSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSGASLISINKLUEHYN RQEEWRRQDVFSLGGLGGLSRLGISLLGAI WAGEWRRCDLFFSLGCGLARLIGGLSRLGISHYK HISKGTLTSTINLATSLARNDMLSLSLDEHYN RQEEWRRQDVFSLGEGLGGLSRLGISLLGAI				i			
826 2176 A 6744 3 5177 SDDLRTGLEQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPERGLTLVRIJPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVEV AFREFINLSESKYCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSGSTCDPL VTFTALAACTRVDSCFTFWFYPSLCVSFQFAH LEFHLCHHLDQLGTAAPQVLQFFVSDRMMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVILFSIFGQMAVS SDVVEKLLDQLGTVAPVNLGQHVHSLNT AIQAWQQNKCFEVELVFSHFVICDTQCTL RFGQVTDTDENILLASLHSQYSWRSHKSFQL LHCIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDQQAVREHFDCLTAKQK LPSYLLENNELTELCVKAKGEDEWSRDVCLE SKAPEYSIVQVPSSNSSITYWCTVVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLERRSLGL SETQIPGKGQEKPLQNIPGVTVTVLTULEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLERRSLGL SETQIPGKGQEKPLQNIPGVTVTVLTULEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLERRSLGL SETQIPGKGQEKPLQNIPGAFQIGITY WANTNTVHKSVAIKLVHNLTSFKWKDGGNG EVYTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIDNTTPYQIFYKPQLSV CNPHSKEYFRVPDSATFSICPGGEQPAMKS SLPCWDLMPDISQSVLDASLLQKQMLGFSPA PGADSSQCWSLPAIVRPERGVSVAVPLGNFR ENGCTRAIVLTYQBHLGYTYLTLSEDPSFRV HINKCPVKMLKENIKQHEVFEVYCKKIPSECS HHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGGFQYYDVD VHQCGTVFTVAPEKKAGPLTNTNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFDLTTHNRAPEKIV TFKMFITQLSLAVFGRGGASALVSGIGGASTSFVK HISKGTLTSTITLALARNMAPILSLLGEHYM AGALFRAGWVOGSLIDASALVSGIGGASTSFVK HISKGTLTSTITLATSLARNMAPILSLEGHYM RQEEWRRQLFESLGCGLSRGGISRGTISFVK HISKGTLTSTITLATSLARNMAPILSLEGHYM RQEEWRRQLFESLGCGLGG	İ			1			· · ·
TEDCPGMMLWRYPEPRGLTL.VRITPVPPNTT EDPDISTADLGDVLOPPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDMLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSTCDPL VTTFALAACRTVDSCTTPWFYPSLCVSFQFAH LEFHLCHHLDQLGTAAPQVLQFYSDRNMPS ELESYMIVSFREPHMYLRQGVVPGSDRNMPS ELESYMIVSFREPHMYLRQGVVVPSDRNMPS ELESYMIVSFREPHMYLRQGVVVFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AQAWQQNKCFEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICEGWGNWRWSEPSSYDHAGTFIRITQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDQGAVCHFIPDCLTAKQ LPSYLLENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSIIYVWCTVLTLEPNS QVQQRMIVFSFLFIMRSHLPDFIIHLERCSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTPQAREE YDPSDCAVFDSTSLKQUATKVHPGGTVNQLD EFYGPEKSLQPIWFYNKKDSDRNEQLSQWDS PMRVLSIWFPYNTLIELLPWALLNISSKW DLWLFBGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTERILGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTFYQIFTXPQLSV CNPHISGKEYRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISGSVLDASLLQKQMLGFSFRA PGADSSQCWSLPAINFREPGSVAVPLGNFR ENGFCTRAIVLTYQEHLGYTVLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYXKLIPSELG HELY HQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDNSQCTQVVFLTGFGYYYDVD VHQGGTVTIVAPEGKARDLIFNINNTPPQIEITNNTRAPEKIV TFKMPTTQLSLAYFDDLTHHKASABLLRLTL DNIFLCVARGAGPLFGEGRYAVALFELLVCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEVKEKCFFIKLCTINNEG KSLCDINEFSFELKPARLYVEDTTVYYKILT DTYLPNSRLAGHSTHLSGGKQVLYRLCTINGA KSLCDINEFSFELKPARLYVEDTTVYYKILT DTYLPNSRLAGHSTHLSGGKQVLYRLCTINGA KSLCDINEFSFELKPARLYVEDTTVYYKILT DTYLPNSRLAGHSTHLSGGKQVLYRLCTINGA KSLCDINEFSFELKPARLYVEDTTVYYKILT DTYLPNSRLAGHSTHLSGGKQVLYBLALAMHY AAGALFRAGWVVGSLDILGSGAALVENALMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AAGALFRAGWVVGSLDILGSGAALVSHALAMHY AGGEWRRQUPESLGECGLARGGLSRLGISLIGAH WANDALLSLAGHAMARLSLLGAHYN RQEEWRRQUPESLGECGLARGGLSRLGISLGAH	[_		L -		L		
TEDCPGMML WRYPERGLIT. VERTIPYPNITI EDPIDISTALD LGDVLO DPCSLEYWDEL (OK VFV AFREFNLSESK VCELQLPDINL VNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSTCDPL VTTFALAACRTVDSCTIPTWYPSLCVSFQFAH LEFHLCHHLDQLGTAAPQVLQFFYDRINMPS ELEFYMIVSREPHMYLR (OWNNGSVCQEIQFE AQADCKLLECRNVTNGSVVKPFSIFGQMAVS SDVVEKLLDCTVVDSVFVNL GQHVVHSLNT AIQAWQQNKCFEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHSSPQL LHCLEGWONWRWSEPFSVDHAGTFIRTIQYR GRTASLUKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQVHYGOQGQAVCHFIPDCLTAKQ LPSYLLENNELTELCVKAKGDEDWSRDVCLE SKAPFYSIVJOVPSSNSSIYVVCTVLTLEPNS QVQQRMIVFSPLFINGRSHLPDPIIHLEKRSLGL SETQIIPGKGGEKPLQNIEPDLVHALITQAREE YDPSDCAVFDSTSLIKQUATKVHPGGTVNQLLD EFYGPEKSLQPIWFYNKKDSDNEDCLSQWDS PMRVKLSIWFYYNTLIELLPWALLNESKW DLWLFBGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTERILGAFPGHQKLCQFGIS SMVQQGIQIIQIEDKTTIINNTFYQIFFXQLSV CNPHISGKEYRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISGSVLDASLLQKQMLGFSPRA PGADSSQCWSLPAINREFFRQSVAVPLGNFE ENGFCTRAIVLTYQBHLGYTVLTLSEDPSPRV IIHNCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDNSQCTQVVFLTGFGYYYDV VHQGGTVFITVAPEGKAGPLJEFRQSVAVPLGNFE ENGFCTRAIVLTYQBHLGYTVLTLSEDPSPRV IIHNCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKAULFSLLLRVEPLDE VTTEWSDAIDNSQCTQVVFLTGGGTYYYDV VHQGGTVFITVAPEGKAGPLJEFRQSVAVPLGNFE ENGFCTRAIVLTYQBHLGYTVLTLSEDPSPRV IIHNCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKAULFSLLLRVEPLDE VTTEWSDAIDNSQCTQVVFLTGGGTYYYDV VHQGGTVFITVAPEGKAGPLJEFRQSVAVPLGNFE ENGFCTRAIVLTYQBFLCYPYLIVKINTAPEKIV TFKMFITQLSLAYFDDLTHHKASABLLRLTL DNIFLCVARGAGPLJEBERYLTNINRAPEKIV TFKMFITQLSLAYFDDLTHHKASABLLRLTL DNIFLCVARGAGPLJEBERYLTNINRAPEKIV TFKMFITQLSLAYFDDLTHHKASABLLRLTL DNIFLCVARGAGPLJEBERYKEKCFFIKLCTILNEG KSILCDNEFFFELKPARLYYEDTTVYYTKLTL DTYLPNSRLAGHSTHLSGGKQVUSCHTSTVFK ASAHTPLSTSVFERGPFFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSAALVSUNGSTGSVK HISKGTLTSITNLATSLARNMRUSLDEHYN RQEEWRRQUVGSLDILGSAALVSGUNGGTTSFVK HISKGTLTSITNLATSLARNMRUSLDEHYN RQEEWRRQUVGSLDILGSAALVSGUNGGTSFVK HISKGTLTSITNLATSLARNMRUSLDEHYN	826	2176	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE
EDPDISTADLGDVLQPCSLE_WDELQKVFV ARFERILSESK VCELQLPDNLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHLEIDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQADCLALECRNVTMYLRQWNNGSVCQEIQFL AQAQCALLECRNVTMYLRQWNNGSVCQEIQFL AQAQCALLECRNVTMAGTURIVLONDTOLL REFQQVIDTEDNLLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRYSALIKVQQLNGVQKQIIICGRQIICSYLSQS SIELKVVQHYIQOBQQAVWREHFDCLTAKQR LHCIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRYSLIENVQLNGVQKQIICGRQIICSYLSQSVL SETQIPPGKGQEKPLQNWCTVLTLEPNS QVQQRMIVFSPLFMRSHLPDPIHLLEKRSLGL SETQIPPGKGQEKPLPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATXVHPGGTVNQLID EFYGPEKSLQPIWPYNKKSDRNNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGGRGVEKPLQNYAGKIIPPHGEAFQIGIY WANTNTYNKSVARLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAPPGHQKLCQFCIS SMVQQGQIQIQIEDKTINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGQPAMKSS SLPCWDLMPDISQSVLDASLLQQMLGFSPA PGADSSQCWSLPAIVREPFRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSVPDCKTKDLJPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFTTVAPEGKAGPILTNTNRAPEKIV TFIKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLPGKAGPILTNTNRAPEKIV TFIKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLPGKAGPILTNTNRAPEKIV TFIKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLPGKAGPILTNTNRAPEKIV TFIKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLPGKAGPILTNTNRAPEKIV TFIKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLPGVKCYPMQVTQH QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCDINFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGRGVLPMQVTQH ARALVNPVKLRKLVQPVNLLVSIHASLKLYI ASDHTPLSFSVERGPIFTTARQLVHALAMHY AAGALFRAGWWVGGLIGGRAGPIGSRLGISTLGBHYN RQEEWBRQLPSSLGEGRGGISRLGGISLGGBILGBEHYN RQEEWBRQCLPSGLGGGRGGISGLGGBILGGBILGGBILGGBILGGBILGGA					1		TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT
AFREFNI.SESK.VCELQLPDINL.VNDQKKLVSS DLWRIVLNSQNIGADDQSSASESGQSTCDPL VTPTALAACTRVDSCTTPWFVPSLCVSFQFAH LEFHILCHILDQLGTAAPQYLOFFYSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVXPFSIFGQMAVS SDVVEKLLDCTVIVDSVYVNLGQHVVHSLNT AIQAWQQNKCFEVELEVSHFVICNDTQETL RFGQVJTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIICGRQIICSYLSQ SIELKVQHYIQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVQVPSSNSSIITYWCTVLTLEPNS QVQQRMIVFSFLFIRMSHLPDFIIHLERSLGL SETQIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPHWPYNKKDSDRNEQLSQWDS PMRVKLSIWKFYVRTLLELLEWALLNESKW DLWLFEGEKVTLQVPAGKIIIPPHQEAFQLGYW WANTNTVHKSVAKLVHNLTSFKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTITINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATTSICPGEQPAMKSS SLPCWDLMFDISGSVLDASLLQKQDMLGFSPA PGADSSQCWSLPAIVRPEPPROSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSRV IHNNCPVKMLIKENIKDIFKFEVYCKKIPSECS HHELYHQISSYPDCKTKDLLFSLLRVEPLDE VTTEWSDADDNSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPLITNTNRAPEKIV TFKMPITQLSLAVFDDLTHIKASAELLRLIL DNIFLCVAPGAGPLPGEPEVAALFELLYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEFYKEKCFIKLCTLNEG KSILCDNFFSFELKPARLYVEDTFVYYWKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVMPYKLRSLIVQPYNLLVSHASLKLYI ASDHTPLSFSVFRGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLIGSPASLVSGIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSTINLATSLARNMDRSLIDEEHYN RQEEWBRQLJESAGLUSLRGAISLLGAI						l	1
DLWRIVLNISSONGADDQSSASSGSGSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHICHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCLLECRNIVTMOSVVYFSIFGOMAVS SDVVEKLLDCTVYDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVELVSHPVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICEGWGNWRWSEPFSVDHAGTFIRTITYY RGRAISLIKVQQLNGYQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYLEINERLFLELVAKAGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIHHLEKRSLGL SETQIPPGKGGEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQLATKVHFGGTVNQLL EFYGPEKSLOPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLELLPWALLNESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAKLVHNLTSFKWKDGGNG EVVTLDEEAFVDTEIRLGAPFOHQKLCQFGIS SMVQQGIQIQIEDKTITINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICFGGEQPAMKSS SLPCWDLMPDISQSVLDASLQKQMLGFSPA PGADSSQCWSLPAIVRFEFPRQSVAVPLGNFR ENGFCTRAIVLTVQEHLGVTVLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSVPDCCTKDLJPSLLLRVEPLDE VTTEWSDADDNSQCTVVRLTGFGTYVYDDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMFTTQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPAGGPLFGEPVAALFELLYCVEIC CGDLQLDNQLYNKSIPHFAVLVCQGEKAEPI QCSKMOSLLSINKELEFKKECFHLCTILNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGRQCVLPMQVTOH ARALVNPVKLRKLUVQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLIGSPASLVSGING VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRSLDEEHYN RQEEVWRQLPSLEGGRGGLSRLGISLGAI	1		ļ	1			1
VTPTALAACTRVDSCFTPWFVPSLCVSFOFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWINGSVCQEIQFL AQADCKLLECRNVTMQSVVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQFVVHSLNT AIQAWQQNKCFEVELLYSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICEGWGNWRWSEPSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQMIGGRQIICSVLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYLENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSITYWCTVLTLEPNS QVQQRMIVFSPIFMSSHLPDPIHLEKRSLGL SETQIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQLLD EFYGPEKSLQPWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAPQIGIY WANTNTVHKSVAIKLVHNLTSPKWCDGGNG EVVTLDEEAFVDTEIRLGAFFGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRYPDSATFSICPGGEQPAMKSS SLPCWDLMPDISGSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFFRQSVAVPLGNFR ENGFCTRAVLTYPGLIGVTLY ILSEDPSPRV IHNRCPVKMLIKENKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLIRVEPLDE VTTEWSDAIDNISQCTQVVPLTGFGYVYDVV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPIPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFIHSAVLVCQGEKAFH QCSKMQSLLISMKELEFYKECFIKLCTILNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AGGALFRAGWVVSGLIGSGLRGGLSRUGSLSEEHYN RQEEWRRQIFESIGGGLRQGLSRLGISLIGAILLELEPTN		į			1	1	1
LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSPREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNYTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVELVPSHFVLONDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIKVQQHXIGGRQGICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPFYSIVQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPPPIIHLEKRSLGL SETQIPGKGQEKPLOHEPDLVHHLTQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVTRTLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIEDKTTIINNTPYQIFVKPQLSV CNPHSGKEYTRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQWSLPAIVRPEFFRQSVAVPLGNFR ENGFCTRAVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELLYNGISSYPDATHFFAVLVCQEKAEPI VTTEWSDAIDNSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPLINTINRAPEKIV TFKMFITQLSLAVFIFFAVLVCQEKAEPI QCSKMQSLLSNKELEEYXEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYJKTLF DTYLFNSRLAGHSTH.SGGKQVLYMQVTOH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVSGLISNKELEEYXEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYJKTLF DTYLFNSRLAGHSTH.SGGKQVLYMQVTOH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVSGLIGSFASLVSGIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSTITNLATSLARNMDRLSLDEEHYN RQEEWRRQLEESLEGGLRQGLSRLGISLLGALLELTY	1	l				ł	
ELEYMIVSFREPHMYLRÖWNNGSVCQEIQFL AQADCKLLECTNYTMQSVVKPFSIFQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVELLVFSIFVICNDTQETL RFGQVDTDENILASLHSHQYSWASSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIGGRQICSSVLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPFYSIVQVPSSNSSITYWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPPPIHLEKRSLGL SETQIPGKGGEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWRPYVRTLIELLPWALLINESKW DLWLFGEKIVLQVPAGKIIIPPPICPAFQIGTY WANTINTHKSVAIKLVINLTSPKWKDGGNG EVVTLDEAFVDTEIRLGAFGHQKLCQFCIS SMYQQGIQIQIEDKTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLIQKQIMLGFSPA PGADSSQCWSLPAJVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIFNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSHFFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTILNEG KSILCDINEFSFELKPARLYVEDTFYYYIKTLF DTYLPNSRLAGISTHLSGGKQVLPMQVTOH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSGYTTSPVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLFESLGGLRQGLSRLGISLLGAL		ŀ	1	İ		[	1
AQADCKLLECRNVTMQSVVKPFSIFGÖMÄVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPVEELVFSHFVICNDTQETL RFGQVDTDENILASLHSHQYSWRSHKSPQL LHICIEGWGNBWSEPFSVDHAGFIRITIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIJYVWCTVLTLEPNS QVQQMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIFGKGGEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQUATKVHPGGTVNQLLD EFYGPEKSLQPIMPYNKKDSDRNEQLSQWD PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFFGGKVLQVPAGKIIIPPNFQEAPQIGIY WANTITVHKSVAKLVHNLTSPF WKDGGNG EVVTLDEEAFVDTEIRLGAPFGHQKLCQFCIS SIMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYRFVPDSATTSICPGGEQPAMKSS SLFCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFFRQVAVPLGNIF ENGFCTRAIVLITYQEHLGVTYLTLSEDPSRV IIHNRCPVKMLIKENIKDPKFEVYCKRIPSECS HHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDADINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPLTNITNAPEKIX TFKKMFITQLSLAVPDDLTHHKASAELLRLTL DNIFLCVAPGAGPL.PGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEFYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHTSILSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSTSVFERGPEITTARQLVHALAMHY AAGALFRAGWVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPSSLGISSLGISLLGISL	ļ	1	į	l	l	ļ	
SDVVEKLLDCTVIVDSVEYNI.GOHVVHSILNI AIQAWQQNKCPEVEL VFSHFVICNDTQETI RFGQVDTDENILASI.HSHQYSWRSHKSRQL LHICEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASI.HKVQQLNGVQKQIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSINQVPSSNSSIIYVWCTV.TLEPIS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGGEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTLKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNTQEAPGIGTY WANTITVHKSVAKLVHNI.TSPK WKDGGNG EVVTLDEAFVDTEILLGAFPGHQKLCQFCIS SMYQQGQIQIIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVREFFRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYL.TLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSYDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPIL.TNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPILPGEEPVAALFELYCVEIC CGDLQLDNQLYNSIPHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHTBLGSGKGVVLPMQVTQH ARALVNPVKLRKLVIQPVNILVSHASLKLYI ASDHTPLSFSVFERGPITTARQLVHALAMHY AAGALFRAGWVOSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPSLGEBLRQGLSSLGISLLGAI		ļ	1	1		ĺ	
AIQAWQONKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICCEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYLLENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQYPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SKAPEYSIVIQYPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIFGKGGEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQLLD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAPQIGIY WANTNTVHKSVAIKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAPFGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVDBSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFRQSVAVPLGNFF ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADDNSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPLLTNTNRAPEKIV TFKMFTTQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTTLNEG KSILCDIMFSFELKPARLYVEDTFVYYKIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVGSLDLIGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSOVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPSSLGGELRQGLSRLGISLLGAI		l		l	İ		
RFĠQUTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVQVPSSNSSIIYWCTVLTLERNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIPGKGGEPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQLATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWFFGEKIVLQVPAGKIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSFK WKDGGNG EVVTLDEAFVDTEIRLGAFPGHQKLCOFCIS SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICFGGEQPAMKSS SILPCWDLMPDISQSVLDASLIQKQMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDADIDNSQGTQVVPLTGFGYVYVVV VHQCGTVFITVAPEGKAGPULTNTNRAPEKIV TF/KMPTTQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPILPGEPVAALFELLYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCTTLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPOAFVSGVSRGTTTSFVK HISKGTILTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQIPESLGFGLRQGLSRLGISLLGGAI	1						SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT
LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVPSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQLATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKAVIKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHISGKEVFYPDSATTSICTGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLAAIVREFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDADINSQGTQVVPLTGFGGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMFITQLSLAVFDDLTHHKASAELLRLTL DNIFTCVAPGAGPLPGEEPVAALFELLYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRGPIFTTARQL VHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI					İ	ł	AIQAWQQNKCPEVEELVFSHFVICNDTQETL
LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVPSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQLATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKAVIKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHISGKEVFYPDSATTSICTGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLAAIVREFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDADINSQGTQVVPLTGFGGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMFITQLSLAVFDDLTHHKASAELLRLTL DNIFTCVAPGAGPLPGEEPVAALFELLYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPYNLLVSHASLKLYI ASDHTPLSFSFELFRGPIFTTARQL VHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	Į.	ł	ł		1		RFGOVDTDENILLASLHSHQYSWRSHKSPQL
GRTASLIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHTPCLTAKQK LPSYLEINNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQMTVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEDDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGQIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGGADSSQCWSLPAIVRPEFPRQSVAVPLGNFFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFIFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNILLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	ĺ	1	ľ	ſ	•	Í	
SIELKVQHYĬGODGQĂVVREHFDĊLTAKOK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVOYPSSNSSIIYVWCTVLTLEFNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGGEKPLQNIEPDLVHHLITQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQLID EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVQVPAGKIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMPITQLSLAVFDDLTHIKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYJKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNILLVSHASLKLYI ASDHTPLSSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPSSLGESLLGISLLGAI			ļ	İ		ļ	1
LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLITLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRILLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATTSICPGGEQPAMKSS SLPCWDLIMPDISQSVLDASILQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHHNCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLIL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKFARLYVEDTFVYYKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRQLPESLGGGLRQGLSRLGISLLGAI	1	1	İ				
SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSELFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAPFGHQKLCQFCIS SMVQQGIQIEDGKTTIINNTPYQIFYKPQLSV CNPHSGREYFRVPDSATTSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNNCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSTINLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGGGLRQGLSRLGISLLGAI	l	1	ł	i .			
QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGI. SETQIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNERCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYYYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TFKMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI			1	ļ	ĺ		1
SETQIPGKGQEKPLQNIEPDLVHHLITQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVARKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYYYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHPAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIOPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSGTTTSFVK HISKGTLTSTINLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	j .	1	ł	1	1	i	<b>1</b>
YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	l		1				
EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	ľ		1	ŀ			
PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTINTVHKSVAIKLVHILTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQMLGFSFA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSFRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	]		i	l			
DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTYHKSVAIKLVHNILTSPK WKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFTLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYYEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYJIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI				ŀ			EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS
WANTNTVHKSVAİKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGİQIQEDKTTİINNTPYQİFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQİMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLİKENIKDİPKFEVYCKKİPSECS IHHELYHQİSSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADİNSQGTQVVPLTGFGYVYVDV VHQCGTVFİTVAPEGKAGPİLTNTNRAPEKİV TF/KMPİTQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEİC CGÖLQLDNQLYNKSNFHFAVLVCQGEKAEPİ QCSKMQSLLİSNKELEEYKEKCFİKLCİTLNEĞ KSİLCDINEFSFELKPARLYVEDTFVYYİKTLF DTYLPNSRLAGHSTHLSGĞKQVLPMQVTQH ARALVNPVKLRKLVİQPVNLLVSİHASLKLYİ ASDHTPLSFSVERGPİFTTARQLVHALAMHY AAGALFRAĞWVQSLÜLĞSPASLVRSİĞNĞ VADFFRLPYEĞLTRGPGAFVSGVSRĞTTSFVK HİSKĞİLTSİTNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLĞEĞLRQĞLSRLĞISLLGAI	i					1	PMRVKLSIWKPYVRTLLIELLPWALLINESKW
WANTNTVHKSVAİKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGİQIQEDKTTİINNTPYQİFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQİMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLİKENIKDİPKFEVYCKKİPSECS IHHELYHQİSSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDADİNSQGTQVVPLTGFGYVYVDV VHQCGTVFİTVAPEGKAGPİLTNTNRAPEKİV TF/KMPİTQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEİC CGÖLQLDNQLYNKSNFHFAVLVCQGEKAEPİ QCSKMQSLLİSNKELEEYKEKCFİKLCİTLNEĞ KSİLCDINEFSFELKPARLYVEDTFVYYİKTLF DTYLPNSRLAGHSTHLSGĞKQVLPMQVTQH ARALVNPVKLRKLVİQPVNLLVSİHASLKLYİ ASDHTPLSFSVERGPİFTTARQLVHALAMHY AAGALFRAĞWVQSLÜLĞSPASLVRSİĞNĞ VADFFRLPYEĞLTRGPGAFVSGVSRĞTTSFVK HİSKĞİLTSİTNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLĞEĞLRQĞLSRLĞISLLGAI	Ĭ.	İ	1	1	[		DLWLFEGEKIVLOVPAGKIIIPPNFQEAFQIGIY
EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPYKLIKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI		1	1	1		ļ	
SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	ļ		1	1			
CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1		1	1	1		
SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IIHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPTTQLSLAVFDDLTH-HKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI			1		l	1	
PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI			1	1	[	1	
ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1		1			1	
IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	ĺ	1	í	1	1	[	
IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	l		1				
VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1		1			
VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI		1	1	1		1	IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE
VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI		1	1		1	1	VTTEWSDAIDINSOGTOVVFLTGFGYVYVDV
TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1	}	1	1	1	
DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1	1	1		1	
CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1	1	1	1		
QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1		1		1	ļ	
KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI			1	1	ľ		
DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI			}				
ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	i	1	1	1	1	1	
ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1	1	!		1	DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH
ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1					
AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI		l	1		1		
VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	1	1	1	Ι΄		1	
HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI	i	}	1	J	J	•	
RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI		1	1	1			
			1	1	I		
AGIVDQPMQNFQKTSEAQASAGHKAKGVISG		1	1	1			
<u></u>	L	1	<u> </u>	<u> </u>	1		AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

GLSQLPKQRHQPSD\\	ntarnic Acid, ycine, H=Histidine, L=Leucine, uragine, P=Proline, ine, S=Serine, yM=Tryptophan, yn, *=Stop codon,
nucl- eotide seq- uence USSN location corresponding to last amino ng to first amino acid residue of peptide sequence sequence sequence sequence location uence uence uence uence uence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location del sequence location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow peptide sequence location del sequence location del sequence location del sequence location del sequence location del sequence location del sequence location del sequence location del sequence location del sequence location del sequence l	ycine, H=Histidine, L=Leucine, aragine, P=Proline, ane, S=Serine, by W=Tryptophan, by vn, *=Stop codon, letion, \=possible
eotide seq- uence USSN location corresponding to last amino acid residue of peptide sequence sequence sequence sequence USSN location corresponding to last amino acid residue of peptide sequence y=Tyrosine, X=Unknow /=possible nucleotide desequence sequence Sequen	L=Leucine, aragine, P=Proline, ane, S=Serine, by W=Tryptophan, by vn, *=Stop codon, letion, \=possible
sequence  09/496 914  orrespondi ng to first amino acid residue of peptide sequence  09/496 914  orrespondi ng to first amino acid residue of peptide sequence  1	aragine, P=Proline, ine, S=Serine, , W=Tryptophan, vn, *=Stop codon, letion, \=possible
uence  914  ng to first amino acid residue of peptide residue of peptide sequence  914  ng to first amino acid of peptide residue of peptide sequence  914  ng to first acid residue of peptide residue of peptide sequence  914  12-  13-  14-  15-  15-  16-  16-  16-  16-  16-  16	ne, S=Serine, , W=Tryptophan, vn, *=Stop codon, letion, \=possible
amino acid residue of peptide residue of peptide sequence y=Tyrosine, X=Unknow /=possible nucleotide de sequence vGKGIMGVFTKPIGO.  Barrino acid of peptide T=Threonine, V=Valine, Y=Tyrosine, X=Unknow /=possible nucleotide de nucleotide insertion vGKGIMGVFTKPIGO.  GLSQLPKQRHQPSD\\	, W=Tryptophan, vn, *=Stop codon, letion, \=possible
residue of peptide peptide sequence y=Tyrosine, X=Unknow peptide sequence nucleotide de pucleotide insertion VGKGIMGVFTKPIGO GLSQLPKQRHQPSD\\	vn, *=Stop codon, letion, \=possible
sequence nucleotide insertion  VGKGIMGVFTKPIGG  GLSQLPKQRHQPSD\\	· •
VGKGIMGVFTKPIGG GLSQLPKQRHQPSD\\	AAELVSQTGYGILHGA
GLSQLPKQRHQPSD\\	AAELVSQTGYGILHGA
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	JHADQAPNSHVKYVW
	LDVVLVRGSGQEHEGC
	DTQQQAFPVTEIDCAQD
	VACDVEVDGVRERLSE
	HLAPSCSSMQIPCPVVA
	DPHFAQVFLSKFTMVK
NKALRKGFP	TO CONTROL OF COMME
827 2177 A 6748 2 1662 FVGAPRRGNPFGSPG	NPGRHQGPCHRPRGTK
	ATGLEMPSSGRALLDSP
	EGEGEPLALGDCFTVN FPHTRLGKLAVVVASY
	CDDANPVDNEYFFDRS
	RLHVMEQLCALSFLOEI
	DRYFRRKELSETLDFKK
	SQGPCPTVRQKLWNIL
	IIFVGVSIINMALMSAEL
SWLDLQLLEILEYVC	ISWFTGEFVLRFLCVRD
	JLPFYITLLVESLSG\SQT
TQEL\ENVGAHCPGC	LRLLRAL\RMLKAWGR
	YEEVGLLLLFLSVGISIF
1 1 1 1 1	TSVPCAWWWATTSMT
)	IVAFMCILSGILVLALPI
1 1 3 1 3 + 1	KEAAVRQREALKKLTK
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	YARSIMEMLRLKGRER
	VCA ATREEL CERVINEY
	AVSAATREELGEPVPFV VRSSPSGNTQSSPKSKQ
	PQLDSKFSNQGKQGGS
	HTPKALPGPGGSMGLK
	RSISADSFDQRDPGTPN
	SQDSQHTPHSMTPSNAT
	EPTPAQKTPAKVVYVFS
	QVETIVSFHIQNISNNK
TERSTAPLNTQISALR	NDPKPLPQQPPAPANQ
DQNSSQNTRLQPTPP	IPAPAPKPAAPPRPLDRE
1 1 1 1 1 1	ASSTPLPPDGTGPNSTPN
	ADPKAPPPPPVSSGEPPT
LGENPDGLSQEQLEH	RERSLQTLRDIQRMLFP
	NPGVLDGPQKKPEGPI
QAMMAQSQSLGKGP	GPRTDVGAPFGPQGHR
DVPFSPDEMVPPSMN	ISQSGTIGPDHLDHMTP
	KRRKPEQVVVQQCSLQ
	RGPPPPYQMTPSEGWAP LPPRGMAPHPNMPGSO
, , , , , , , , ,	EGPNVPNPASRPGLSGV PPGQGIFSGPGRGERFP
	AEKQLGLPPGMAMEGIR
	MEPGNNPIFPRIPVEGP
	IGPGRELEFGMVPSGM
	QMIPQKMREAGAGPEE
	QQKMVPLPFGEHPQQE
	PGSNSGLRNLREPIGPDO
	PSSNPTSLNTAPPVQRG
	HSPGINPLKSPTMHQVQ
SPMLGSPSGNLKSPO'	TPSQLAGMLAGPAAAA
SIKSPPVLGSAAASPV	HLKSPSLPAPSPGWTSS
PEPPLQSPGIPPNHKA	PLTMASPAMLGNVESG
GPPPPTASQPASVNIP	G\SLPSSTPYTMPPEPTL
SQNPLSIMMSR\MSK	FAM/PS\SNPGYNHDAI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KTVASSDDDSPPARSPNLPSMNNMPGMGINT QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP HNGPSGGQGSFPGGMGFPGGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ YFPRGEVPGRKQPQGPGPGFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHHRMMSPAQST MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
	2155	<u> </u>	-	420		HPGPVGSPGMMMSMQGMMGP\NRTS
829	2179	A	6797	433	3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG GEREPRDRPSERPRLV
830	2180	A	6800	3	1911	LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\ RELRLRATRA\FVAAR\SFVQGLGVAS\DVVR KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDITLTAKVIQGCGNPKVNPQGPGP EEKRRGKLAPRERPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRKPLFY LVNVIAPCILITLLAIFVFYJLPDAGEKMGLSIF ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWQFVAMVVDRLFLW TFIIFTSVGTLVIFLDATYHLPPPDPFP
832	2182	Α	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYW\INPTL\IS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV IMPCNWWVARMLGRV  EAEGEQVCGAKCCGDAPHVENREEETARIGP
						GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVVN LYSLLERINPDHSFPVSSHCLRAAAFVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6862	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG PDEILL DI BATGACGAVAASTTGSSCDC/SGT
		•				RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP
837	2187	A	6863	2	1615	VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEFKVL LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
]	}	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
,				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide	Joquonoo	/=possible nucleotide deletion, \=possible
						nucleotide insertion
<u></u>	<u> </u>		<u> </u>	sequence	<u> </u>	
[		ļ		ļ		PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR
i	1	l		ŀ	ł	APVSAYQYALANGDVWKVHEVPDYSMAYG
1	Ĭ	ĺ				NPGVADATPPWSSYKEQSPQTLLELKRQRAA
	1				j	AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI
	1				1	GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
						EEMEEKVHGCCRIS
838	2188	A	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS
1 050	-100	1	1 0000	323.	'**	LRPRKLDFFRSEKELNHLAVDEASGVVYLGA
ł	1		ł		1	VNALYQLDAKLQLEQQVATGPVLDNKKCTP
		l				PIEASOCHEAEMTDNVNQLLLVDPPRKRLVE
					1	
		ļ	1	Į.		CGQLLKGI\CALRALSNISLRLFYEDGSGEKSF
		1		1	1	VASNDEGVATVGLVSSTGPGGDRVLFVGKG
{	1	[	ĺ	1	1	NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT
		1		1	1	YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD
	1	1		1	1	KHPARNRTLLARMCREDPNYYSYLEMDLQC
1		1		1		RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF
i	1	1	1			SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN
1	ļ		]	j	j	ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK
1	İ		1			SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN
1		1			1	LTAVTVAAENNHTVAFLGTSDGRILKVYLTP
		1	Į.	İ		DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY
ł	ļ		1	1		AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY
		Ì	1	1	1 .	1
	ľ		-			CGWCVVEGRCTRKAECPRAEEASHWLWSRS
	ļ					KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA
1			i	1		LSEEDELLCLFGESPPHPARVEGEAVICNSPSS
		1	ŀ			IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF
	.	1	ļ		}	YDCRQAMSLEENLPCISCVSNRWTCQWDLR
1			1		1	YHECREASPNPEDGIVRAHMEDSCPQFLGPSP
	1	į.		1		LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD
	1	1			1	LLKFMEPVTMQESGTFAFRTPKLSHDANETL
				1		PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC
1						SLCRAANPDYRCAWCGGQSRCVYEALCNTT
j	1	1		ļ	ļ	SECREATION TROPET GPLGGGIRITILG SNLGVQ
	1	i				1
	Į.	1				AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA
		1				AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP
1	1					KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED
1	1	İ	1	ļ		VRVTLNGVPCKVTKFGAQLQCVTGPQATRG
1	1	Į.	1	}		QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE
	1			Ì		PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP
1	1	1				LQSWQPPREAESLQPMTVVGTDYVFHNDTK
1		1	-			VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT
				i		EAGAFEYVPDPTFENFTGGVKKOVNKLIRAR
	1	1	İ	1		GTNLNKAMTLQEAEAFVGAERCTMKTLTET
1	1	1	1	ł	1	DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF
1		1	1			
		1				GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM
		1	I	ŀ		VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG
1		1	1	1	1	LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG
	1	i	-	l	1	IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL
1		1.	1	l		DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT
		1				L'ENQPEFSARAKVYFASLLTVALHGKLEYYT
	1 .		1	1	1	DIMHTLFLELLEQYVVAKNPKLMLRRSETVV
		1	1	1		ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI
}		1	1	1		KHQVEKGPVDAVQKKAKYTLNDTGLLGDD
}		1	1	1		VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ
i		1	1	}		
		1	1	1	1	VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG
		1	1	I		STAQILSDLDLTSQREGRWKRVNTLMHYNVR
	1		1	1	1	DGATLILSKVGVSQQPEDSQQDLPGERHALL
l	1	1	1		1	EEENRVWHLVRPTDEVDEGKSKRGSVKEKE
ļ	1	1	1	1	1	RTKAITEIYLTRLLSVKGTLQQFVDNFFQSVL
1	1	1		I	1.	APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI
L	<del></del>	т—		<del></del>		

SEQ ID NO: of NO: of nucl- nucl- eotide seq- uence   SEQ ID NO: of nucl- eotide   SEQ ID NE- ID NO: of nucl- eotide   Seq ID NO: of nucleotide   Seq ID NE- ID NO: of nucleotide   D-Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine Corresponding to last amino acid residue of peptide sequence   Seq ID NE- ID NO: of nucleotide   D-Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine Corresponding to last amino acid residue of peptide sequence   Seq ID NE- ID NO: of nucleotide   D-Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine Corresponding to last amino acid residue of peptide sequence   Seq ID NE-	HVHEVV PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
nucleotide sequence  in location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location corresponding to last amino acid residue of peptide sequence  in nucleotide location (p=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, P=possible nucleotide deletion, V=Possible nucleotide insertion  In N=M=N=N=N=N=N=N=N=N=N=N=N=N=N=N=N=N=N=N	IVHEVV PSNKLL 'SDQDM 'YTQKY AAALE LPGQT DPVSLH
eotide sequence  USSN 09/496 914   location corresponding uence   914   possible nucleotide deletion, \= possible nucleotide deletion, \= possible nucleotide deletion, \= possible nucleotide deletion, \= possible nucleotide deletion, \= possible nucleotide insertion  HIWKTNSLPLRFWVNILKNPHFIFDVF DASLSVIAQTFMDACTRTEHKLSRDS YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL/NKVTDL  839 2189 A 6872 1   1485   RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	IVHEVV PSNKLL 'SDQDM 'YTQKY AAALE LPGQT DPVSLH
sequence  uence  09/496 914 914 914 914 914 914 914 914 914 914	IVHEVV PSNKLL SDQDM IYTQKY AAALE LPGQT DPVSLH
uence  914  ng to first amino acid residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  914  125	IVHEVV PSNKLL SDQDM IYTQKY AAALE LPGQT DPVSLH
amino acid residue of peptide sequence	IVHEVV PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
residue of peptide sequence  residue of peptide sequence  y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  HIWKTNSLPLRFWVNILKNPHFIFDVF DASLSVIAQTFMDACTRTEHKLSRDS YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	IVHEVV PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion  HIWKTNSLPLRFWVNILKNPHFIFDVF DASLSVIAQTFMDACTRTEHKLSRDS: YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	IVHEVV PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
sequence nucleotide insertion  HIWKTNSLPLRFWVNILKNPHFIFDVF DASLSVIAQTFMDACTRTEHKLSRDS: YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
HIWKTNSLPLRFWVNILKNPHFIFDVF DASLSVIAQTFMDACTRTEHKLSRDS YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
DASLSVIAQTFMDACTRTEHKLSRDS: YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	PSNKLL SDQDM YTQKY AAALE LPGQT DPVSLH
YAKEISTYKKMVEDYYKGIRQMVQV NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	SDQDM YTQKY AAALE LPGQT DPVSLH
NTHLAEISRAHTDSLNTLVALHQLYQ YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	YTQKY AAALE LPGQT DPVSLH
YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	AAALE LPGQT DPVSLH
YDEIINALEEDPAAQKMQLAFRLQQL NKVTDL  839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	AAALE LPGQT DPVSLH
839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	LPGQT DPVSLH
839 2189 A 6872 1 1485 RARRLALQCHVCVCALTPGEQSGRRI WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	DPVSLH
WLMFSCFCFSLQDNSFSSTTVTECDEI EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	DPVSLH
EDQTDCSSLRDENNKENYPDAGALVI SWEPQQQNVEATVLVDSVLRPSMGN	
SWEPQQQNVEATVLVDSVLRPSMGN	CCLIADD
KSIFKAESGRSHGESQETEHVVSSQSE	
GTPAHESPQNNAFKCQET/VRL\QPRII	
SPKDAFETR\QDLNEEEAAQVHGVKD	
TQSVLA\DGTDSADPSPVHKDGQNEA	
DLHSVGTSRLLL/YHITDGDNPTAVRF	
SGQSQRFNLDPESAPSPPSTQQFMMPI	RSSSRC
SCGDGKEPQTITQLTKHIQSLKRKIRK	
QEKKYRPSHGDKTSNPEVLKWMNDL	AKGRK
QLKELKLKLSEEQGSAPKGPPRNLLCI	
RENGKPEAAGPEPSSSGEETPDAALTO	•
EQLPPQEDSKVTKQDKNLIKPLYDRY	
STPSLIPTIVSQDTCMLLLCTDV	
840 2190 A 6873 2 2054 FFRFYFSFIRLFAMSLADLTKTNIDEHI	FEGVAI
ENNRRSAACKRSPGTGDFSRNSNASN	
SRSQCSCGSLSSQYDYSEDFLCDCSEK	
YLKQPVVKEKEKKKYNVSKISQSKGQ	ATIMAN.
EVELTURAL OF PAICOUR MADERAL	YTIDII O SVEIDA
EKKHTWNASLFNSQIHMIAQRRDAM	AHKILS
ARLHKIKGLKNELADMHHKLEAILTE	NQFLK
QLQLRHLKAIGKYENSQNNLPQIMAK	HQNEV
KNLRQLLRKSQEKERTLSRKLRETDS	
KTLAAQTATKTLQVEVKHLQQKLKE	
EIKNIYSHRILKNLHDTEDYPKVSSTK	SVQAD
	KKATG
NIDHKEKSTEINHEIPHCVNKLPKQED	SKRKY
EDLSGEEKHLEVQILLENTGRQKDKK	
KNIFVKEEQELPPKIIEVIHPERÈSNQE	
EKFKRSMQRNGVDDT\LGKGTAPYTK	
RRHYSFTEATENLHHGLPASGGPANA	
YSHSTGKHLSNREEMELEHS\DSGYEF	
SRIKVKDTTFRDKKSSLMEELFGSGY	
QSSPGVAKGSEEPLQSKESHPLPPSQA	
FGDSKVTVVNSIKPSSPTEGKRKIII	OIOIM
841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYI	רוז דור
NAPAPGTPAASGWQPPTYHSGRAFSA	
SRRGYSSHHGPSWRKKYSLVNRPPGP	
DHAVRPLHGARGGQPPVPQQHVLER(	ŲVQLS
QGQNVVIKVKPPSKSGSASASGAQRG	
DTPWSDQRPREGEGEPPRGQLQPSRP	
TCSVEDPLLVCQKEPGKPRMVKSVGS	
	GVALG
RKLGSHSVASCAPQLLGDRRVDAGH	
SGSVGGPARPASGPRQAREASLVVTC	
RKNNYKWVAASSKSPRVARRALSPR'	VAAFN
VCKASAGMANKVEKPQLIADPEPKPR	
SKPGSAPSKYKWKASSPSASSSSSFRW	
SKDHASQLSPVLSRSPSGD\RPALAHS(	
SKDHASQLSFVLSRSFSGDRFALAHS GETPLSAYKVKTRTKIIRRRGSTSLPGI	
TSPAATAKSHLSLRRRQALRGKSSPVI	rkk 1 LN

0E0 TO 1	CEO IN	14-4	<u> </u>	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noa	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496	correspondi	to last amino	M=Methiorine, N=Asparagine, P=Proline,
seq-	ucncc	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			'''	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ	}	peptide	Sequence	/=possible nucleotide deletion, \=possible
			ĺ	1		nucleotide insertion
		<b> </b>		sequence		KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA
						VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS
		l				·
		İ				LPSWRARRLSLSRSLVLNRLRPVASGGGKAQ
		(	[			PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG
į						QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ
		ł				RSLAIIRQARQRREKRKEYCMYYNRFGRCNR
			1			GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC
						PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV
		ſ				YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC
				i		PDFARRGACPRGAQCQLLHRTQKRHSRRAAT
		1	}	l		SPAPGPSDATARSRVSASHGPRKPSASQRPTR
			l		l	QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS
				•		SSSSSSSPPASLDHE\APSLQEAALAAACSNR
		İ		Ì	i	LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS
						GKPLHIKPRL
842	2192	A	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH
						RNKFKVINVDDDGNELGSGIMELTDTELILYT
						RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC
		· .				QTGQGIFAFKCARAEELFNMLQEIMQNNSIN
		]				VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA
			ŀ			QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL
			İ			PSVGEESTHPLLVAEEQVHTYVNTTGVQEER
		1				KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR
			1	1		DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE
		ļ		]	J	QLGRDQVSGSGANNTEWDTGYDSDERRDAP
			1			SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD
						TQNINNSAQRRTALLNYENLPSLPPVWEARK
			l			LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV
				ĺ		NTENVTVPASAHKIEYSRRRDCTPTVFNFDIR
				İ		RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT
		[		ſ		TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL
		l	ł			PRDDGTSR\KTRHNST\DLPL
843	2193	A	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA
						YTTACVLTTAAVQLELITPFQLYFNPELIFKHF
				i		OIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM
		ĺ	ĺ	ľ		LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS
						L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP
						SOLPGPLSALGAHGIFLVVGELNHCGPFGYCS
		1	1		1	WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL
U-1-	7.77	1	3,23		1	AHCNLRLPGSSNSPASASQVAGITGVCHHAR
		l		1	1	LIFVFSVETGFLHAGQAGLELLTSGDPPASAS
		1		1	1	QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA
		1		1	1	LM
945	2105	_	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS
845	2195	A	עכעט ן	1000	317	
			}	J	J	TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE
		1			1	GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT
					1	LPMEKLAASTEPQGPRPVLGRESVQVPDDQD
	]	ļ.	1			FRSFRSECEAEVGWNLTYSRAGVSVWVQAV
		1				EMDRTLHKIKCRMECCDVPAETLYDVLHDIE
				1		YRKKWDSNVIETFDIARLTVNADVGYYSWR
		1	1			CPKPLKNRDVITLRSWLPMGADYIIMNYSVK
		1		1		HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT
		1	1	1		YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK
		1				
	i					MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP
						LPSLALS\ELSVQHADS\LENIDESAV\AESREE
				ſ		
			:			LPSLALS\ELSVQHADS\LENIDESAV\AESREE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
846	2196	Α	6944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE
		ĺ			ĺ	ELTILGETQEEEDEILPRKDYESLDYDRCINDP
						YLEVLETMDNKKGRRYEAVKWMVVFAIGV
		1				CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS
		l				QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE
						AGSGITEGKCYLYARQVPGLVRLPTLLWKAL
						GVLLTVAAMLLI\GLGSPMHSGSVVGAGLPQ
		1	}			FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG
			1			AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL
					Į	TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV
						VMGVIGGLLGATFNCLNKRLAKYRMRNVHP
		ļ			·	KPKLVRVLESLLVSLVTTVVVFVASMVLGEC
		l				RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP
						NDTYNDMATLFFNPQESAILQLFHQDGTFSPV
		}				TLALFFVLYFLLACWTYGISVPSGLFVPSLLC
			1			GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA
		ļ				AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT
		1	]			LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW
						ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV
			]			SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS
						NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL
					·	RNMCDEHIASEEPAEKEDLLQQMLERRYTPY
		1				PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ
						LVTLLVRGVCYSESQSSASQPRLSYAEMAED
			'			YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF
•						TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
847	2197	Α-	6951	3	1994	IVGIITRHNLTYEFLQARLRQHYQTI NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER
017	2177	' `	0,51	3	1774	LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK
						VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI
•		1			·	SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST
		ľ	1			VGKRKIDQEGRVFQEKWERAYFFVEVONIST
						CLICKRSMSVSKEYNLRRHYOTNHSKHYDOY
		ł				MERMRDEKLHELKKGLRKYLLGLSDTECPE
						QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR
		1				EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE
			]			NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK
		}		•		NFCINWSKLVSVASTGTPPMVDANNGLVTKL
						KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM
			1			DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL
						DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI
						DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM
		1				HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL
		[		!		WETHLTRNNLAHFPTLKLVSRNESDGLNYIP
	I					KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI
	ľ	ı	i		!	DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ
	İ	1			+	
i				j		I LESIMKI SKIK II SIII KIISISMIISMI ESKI
848	2198	A	6985	3	289	LFSIMKLSKTKYCSQLKDSQWDSVLHIAT SVOYI PGRPTRTHASTDAPI MI KETPI PSKTK
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C
848	2198	A	6985		289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM
				963		SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIOM
						SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP
						SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVAVVDIAYACNTVPRMLVNLLHP
						SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT
						SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA
						SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPITVC CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	"""	ļ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
l deniec			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide	Soquenee	/=possible nucleotide deletion, \=possible
Ì		l		sequence		nucleotide insertion
<b> </b>		<del> </del>	┼	sequence	<del></del>	GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS
		ł				LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL
850	2200	A	7001	1	1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
830	2200	Α	1 7001	1	.1011	DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG
1			1	į		l · · · · · · · · · · · · · · · ·
	1					KVARRRVGATWLLHLAVADLLCCLSLPILAV
		1	ļ	ŀ		PIARGGHWPYGAVGCRALPSIILLTMYASVLL
1		1	1			LAALSADLCFLALGPAW\CLRFS/GACGVQVA
ł		1	ł	1	ł	CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ
	!					CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA
1	1	ļ	1	1		SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
i .		1	}	ļ		YHLLGLVLTVAAPNSALLARALRAEPLIVGL
			1			ALAHSCLNPMLFLYFGRAQLRRSLPAACHW
L	<u> </u>	<u> </u>	1			ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI
	}	l	1			SRGVLVCDECCSVHRSLGRHISIVKHLRHSA
1	ŀ	ł	1	ł		WPPTLLQMVHTLASNGANSIWEHSLLDPAQV
ì		ł	i	}		QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF
	1		l .	İ	1	VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE
		ļ		1		TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG
1		ì			1	QTLQAELLVVYGADPGSPDVNGRTPIDYARQ
1	ĺ	ĺ	1	(		AGHHELAERLVECQYELTDRLAFYLCGRKPD
i			I			HKNGHYIIPQMADSLDLSELAKAAKKKLQAL
		l				SNRLFEELAMDVYDEVDRRENDAVWLATQN
İ	ĺ	1			i	HSTLVTERSAVPFLPVNPEYSATRNQGRQKL
	ļ	ļ	j		]	ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD
	ľ			]	İ	NLELSLRSQSDLDDQHDYDSVASDEDTDQEP
		1	l .	1		LRSTGATRSNRARSMDSSDLSDGAVTLQEYL
	ļ		i	i.		ELKKALATSEAKVQQLMKVNSSLSDELRRLQ
1	1	ŧ	1			REIHKLQAENLQLRQPPGPVPTPPLPSERAEH
1	ł	į	1	ł		TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG
		ŀ	1		1	PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL
			:	i		YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK
J		1	j	ļ.	<b>,</b>	LSRHGSGADSDYENTQSGDPLLGLEGKRFLE
						LGKEEDFHPELESLDGDLDPGLPSTEDVILKT
			1			EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA
	1			1		VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ
				l	1	SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
1		1	1	1	ł	KAAKQLVTITTREKKQ
8.52	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL
		1	1	I	l - · ·	TVKGLLKPSFSPRNYKALSEVOGWKORMAA
1		1		1	1	KELARONMDLGFKLLKKLAFYNPGRNIFLSP
		1		l		LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP
1	1	ľ		[	1	EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID
1	1			I		QRLQPQRKFLEDAKNFYSAETILTNFQNLEM
1		l		1	l	ACKQINDFI/ESKTHGKINNLIENIDPGTVMLL
}	1	1	1	1	ì	ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS
				1	1	VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK
	}			1		NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL
1		1	]		1	
]	1	1	1			LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS
1		1	1		l	KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM
				]	İ	DERGTEGAAGTGAQTLPMETPLVVKIDKPYL
	<u> </u>	<u> </u>	<u> </u>	<u> </u>		LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV
		1			-	ARHVAAGAGHENKHGGSRRFPAGVAPRRAM
1	1.	l .	1.	İ		ANVSKKVSWSGRDRDDEEAAPLLRRTARPG
		1	1			GGTPLLNGAGPGAARQSPRSALFRVGHMSSV
	1	ì		1		ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY
1	1		1			ESLDYDNSENQLFLEEERRINHTAFRTVEIKR
	1				1	WVICALIGILTGLVACFIDIVVENLAGLKYRVI
1	1	1	1	1		KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV
					<u> </u>	

NOC of Noc of nucleotide older of peptide of the contract of t	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence wence unner under the corresponding of the stamp of the sequence peptide sequence peptide sequence sequence peptide sequence se		1 ~	1				
uence where where the corresponding to the state of peptide amino acid residue of peptide sequence peptide sequence peptide sequence where the correction of	nucl-	peptide				location	F=Phenylalanine, G=Glycine, H=Histidine,
uence    914   ng to first anino acid residue of peptide peptide sequence   February   F		seq-	1				
anino acid residue of peptide sequence peptide sequence peptide sequence se		uence		1			
residue of peptide sequence   Y=Tyrosine, X=Unknown, %=Sio qodon, %possible nucleotide deletion, wpossible nucleotide deletion, wpossible nucleotide deletion, wpossible nucleotide deletion, wpossible nucleotide insertion   IVGSVIVAFIEPVAGOSGIPOIKCFLNGVKP   VVRI.KTI.VIKVSGUI.SVVGGLAVGKSGFW   VVRI.KTI.VIKVSGUI.SVVGGLAVGKSGFW   HSGSVIVAFIEPVAGOSGIPOIKCFLNGVKP   VVRI.KTI.VIKVSGUI.SVVGGLAVGKSGFW   HSGSVIVAGISQOTSIS.KRIDERVESVERDI   KRDEVSAGAAAGVSAAFAAPVGGVILSIVAGASGAAAGVSAAFAPVGGVILSIVAGASGAAAGVSAAFAPVGGVILSIVAAGVGVIGAVAAAVALVIAVSIVATULISIPACVGGVIGAVAAAVALVIAVSIVATULISIPACVGGVIGAVAAAVALVAAVALVIAVSIRDAGVAAGVGGVIGAVAAAVAAVALVAAVALVIAVSIRDAGVAAGVGGVIGAVAAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAV	uence	}	ł	914			
Pepstide   Requence							
sequence   nucleotide insertion			<u> </u>			sequence	
I.VGSVIVAPIEPYAAGSGIPJIKCFLINGVERP   VVNIKTI.VIKVSQUI.SVVGGLAAVGKEGPM   HSGSVIAAGISQGRSTSI.KRDFKIFEYFRRDI     KRDFVSAGAAAGVAVAKSPERM   HSGSVIAAGISQGRSTSI.KRDFKIFEYFRRDI     KRDFVSAGAAAGVAVASAFARPYGVIESVEI     ASFWIQELTWRIFFASMISTFTLNFVLSIYH     NMWDLSSPGLINFGRFDSSKMAYTHEIPYY     AMGVUSGVIGAVVAAVTATVAFVLIYSSRDCQJ     QGGSMSYPLQLEADGEYNSMAAAFENTP     SVVSLHIPPPGSYNPLTLGLFTI.VYPTI.ACV     YGLTVSAGVFPELLIGAAWGRIFGSILSSYL     AAIWAPPGKYALMGAAQJGGIVRMTLSI.     VIMMHQLQSVPFLHWEAPVTSHSLTARE     MSTPVTCLRREKKVGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVIVDVLSDTASHNIN     PVVEHADDTOPARLQGLIRSQLIVLLKIKK     VERSILGLVQRAKGVITDGLIRSQLIVLLKIKK     VERSILGLVQRAKGLESILAQTGPKQAQTABGI     AGAQQPQCURAVILDGILLIAGGIASP     LEQUVYELQAGVACLREIKQRCREATRSI.     RELLQVRAKQVQLGSEIRQLQOCRAPPAAQ     EKEPPESGLONQMADLDKIL YEVREALT     RRRQVQQEARRGASQCAGGIRLAKLTDILL     QEQQREVACQALQKNQEDSRRVDLEVA     KELLQVRAKYQLQGSEIRQLQOCRAPPAAQ     EKEPPESGLONQMADLDKIL YEVREALT     RRRQVQQEARRGASQCAGGIRLAKLTDILL     QEQQRAVANSVARHGFILAGGIFGDY     HIWHYAATGTILVLLAADAQQGKVFTINT     AVRIPGGRAVANSVARHGFILAGGIFGDY     HIWHYAATGTILVLLAADAQQGKVFTINT     AVRIPGGRAVANSVARHGFILAGGIFGDY     FURNHANDATTINGVIRCHSPRARSIGLINP     HIHYASAWGFEDGGRTYDGPARLAEAFF     GVSQGRGGIGSFFWASGNGGRENDSCC     GYTNSTYTLSISSATGFONPWYSEACSSTL     TYSSONQNEKQIVTTDLRQKCTESHTGTS.     APLAAGIGLTLEANNITVERNDOPILTPKT     GKRLEVEKTYTACLGEPNHITRLEHAQAR     LSYMRRGDLAHLVSPMGTRETLLAAPPID     SADGFNDWAFMTTHSWDEDPSGEWVLEIE     TSEANNYGTILKFTULVTYTTELVQLRSG     GKRLEVEKTYTACLGEPNHTTRLEHAQAR     LSYMRRGDLAHLVSPMGTRETLLAAPPID     SADGFNDWAFMTTHSWDEDPSGEWVLEIE     TSEANNYGTILKFTULVTYTTELVQLRSG     FRGYKVYTIMDRGLISYKGLPFAWQECFOR     GKRLEVEKTYTACLGEPNHTTRLEHAQBECT     GKRLEVEKTYTACLGEPNHTTRLEHAQBECT     GR							
VVRLKTLVIKVSGVILSVVGGLAVGKEGFM HSGSVLAGISQGRETSLKRDKFEFFEYRRDI KRDFVSAGAAAGVSAAFAPVGGVLFSLE ASFWNOPLTWRIFFSAMISTETLINVLSTYM NMWDLSSFGLNFGGFDSKMAYTHEIPUY AMGVVGGVLGAVFNALNYWLTMFRRYTH PCLQVIEAVLVAAVTATVAFVLTYSSRDCOJ QGGSMSYPJCLFCADGEVNSMAAAFFNTFP SVSLFHDPPGSYMPLTLIGLFTLYFSTACV YGLTVSAGVFPSLLIGAAWGRLFGISLSYL' AAIWAPPGKYYAMGAAQLGGVFWMTLSI UNMÆATSNVTYGFPIMLYLMTAKIVGDVY GLYDMHQLQSVYPEHWEAPVTSHSLTARE MSTPVTCLRREKVGVIVDVLSDTASNHNN PVVEHADDTOPARLGGLRSQLVLLKHK VERSIN,GLVQRRLRLKDFRDAYPRFPFIGSL VQBERCTMDLSEFMMPSPYTVGASLP VFKLFRALGLRHLVVVNNRNQVVGLVTRK LARYRLGKRGLESLSAGTOFRAQATAGG AGAAQQPCQLRAVTLEDLGLLLAGGLASP LSLEELSEYSESHFTSTASVPBQDTAKHWI LEQWVVELQAEVACLREIKQRCRATRSL RELLQVRAKVQLQGSELRGLQGCARPAAQ, EKEAPFSGLQNQMALDKRLVSVREAT RRRQVQGGAERGAAGGAGGAGLAAKITDLL QEEQGREVACGALGKNQEDSSRYVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPFMELI WLLWVVAATGTLVLLAADAQGGKVFTNT AVRIPGGPAVANSVARKHGFLALGGIFGDY HFWHRGVTKRSLSPHRYRHSRLQREPOVOV LOQVAKRTKRDVQEPDTMCPPQOWYLS VTQRDLMYKAAWAQQYTGHGTVVSLDAD EKNPPDLAGNYDQFASSPONADDPDPOPR TQMDDNRHGTRCAGEVAAVANNQVCOVC AYNARIGGOVRMLDGREVTDDAYBARSCLINP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFWWASCAGSTL TTYSSGNQNEKGIVTDDARCASSTLT TTYSSGNQNEKGVTTDLRQKCTESHTGTS, APLAAGIGALTLEAKNENTWRDMORIL-VVQ KPAHLNANDWATNGVGRKVSHSYGYGLL LSYNRRGDLAHILVLARPHD SADGFNDWAFMTTHSWDEDFSGEWVLEUF GKRLEVRKTVTACLGEPNHITRLEHAQARI LSYNRRGDLAHILVLARPHD SADGFNUDMFMTTHSWDEDFSGEWVLEUF GKRLEVRKTVTACLGEPNHITRLEHAQARI LSYNRRGDLAHLLVLARPHD SADGFNUDMFMTTHSWDEDFSGEWVLEUF TSRANNYGTLKFTLVLLYGTABEGLPYPPES GCKTLTSSQACVCEEGFSHGKSCVGHCG GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCOGPALTDCLSCPSHASLDPVETCSR QSRESPPQQQPRLPPPVERAGRIRAGLIFE TSRANNYGTLKFTLVLLYGTABEGLPYPPES GCKTLTSSGACVCEEGFSHGKSCVGHCG GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCOGPALTDCLSCPSHASLDPVETCSR QSRESPPQQQPRLPPPVERAGRIRAGHES FERGREGRETATFRUDVETTRASVCAPCHA CATCOGPALTDCLSCPSHASLDPVETCSR GSRESPPQQQPRLPPPVERAGRIRAGLIFE FROMVYTIMDRGLISYKGIPPEAWGEGEF FROMVYTIMDRGLISYKGIPPEAWGEGEF FROMVYTIMDRGLISYKGIPPEAWGEGEF SEEDEGREGRETATFRUDVETTRASVCAPCHA CATCOGPALTPRAEALPGAFFAMAQHFES				1		f	
KRDFVSAGAAAGVSAAFGAPVGGVLFSISH   ASFWORDTURENFASMISTTINPV.SIYH    AMMVDLSSPQLINTGRFDSEKMAYTHEIPTINPV.SIYH    AMMVDLSSPQLINTGRFDSEKMAYTHEIPTINPV.SIYH    AMMVDLSSPQLINTGRFDSEKMAYTHEIPTINPV.SIYH    AMMVDLSSPQLINTGRFDSEKMAYTHEIPTINPV.SIYH    AMMVDLSSPQLINTGRFDSEKMAYTHEIPTINPV.SIYH    PCLQVIEAVLVAAVTATVAFVLIYSSRDCQI    QGSMSYPLQLFCADGFYNSIMAAAFFNTP    SVVSLFIDPTISSISIAAWGRIFGISSSYIT    AAIWADPGKYALMGAAAQLGGIVRMTLSI    VIMMEATSNYTYGFPIM.VLMTAKIVOODIV   GLYDMHIDQLGSVFFLHWEAPVTSISITARE   MSTPVTCLRREEVGVIDVU.SDTASNHINK    PVVEHADDTOPARLQGLLRSVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPDIVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPDIVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPDIVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPDIVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPDIVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPLOVELSIL   VYKLFRALGLRHLVVVDNRNQVGLVILLKHK    VERSILGLVQRRIRLKDFRDAYPRFPLOVELSIL   VYKLFRALGLRHLVVVDNRNQVGLVILLKHK    VERSILGLVQRRIRLSELSILAQFGAQTAFGRIAGATAGGI   AGAAQOPCOLRAVILEDIGILLAGGLASP    LSELSISEYESSIPTTISSYSEPTTYPOPCASIL   VYKLFRALGLRHLVVVQETIPDTAKATATSI   RELLQVYARVQLOGERERICAGOPCATATSI   RERLQVYARVQLOGERROATAGATAGGI   AGAAQOPCOLRAVILEDIGILLAGGLASP    LSELSISEYESSIPTTISSYSERIPTISSYCHILLTONING TOOLOG TOO							VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI
ASFWNOFLTWRIFASMISTITLNFVLSIYM NMWDLSSPGLINGGRDESKATHEIPW AMGVVGGVLGAVFNALNYWLINGFRRYTH PCLQVIEAVLVAAVTATVAFVLTYSSRDCOJ QGGSMSYPLQLFCADGGYNSMAAAFFNTP; SVVSLFHIDPGSYNRLIGLGFLGISLSYL' AAIWADPGKYALMGAAAQLGGIVRMTLS; LOWERTSNVTYGFPMLVLMMTAKIVODVI GLYDMHIQLQSVPFLHWEAPVTSHSLTARE MSTPYTCLRREKVGVIVDVLATSKY VERSILGLAW GRUFMELSTANE MSTPYTCLRREKVGVIVDVLATSKY VERSILGLAW GRUFMLSEFANDESPTTVPQEASLY VFKLFRALGLRHLVYDNROVVCLATRKY VERSILGLVGRRYRPPTQSIS VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDNROVVCLATRKY VFKLFRALGLRHLVYDRORGERAPAQ EKCAFFESGLQNOMQALDKRLVEVREALTH RRRQVQQEARRYAGCLGERLAKLTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA M  MEPPEPPDQAACTAGARGAGRAGGAGGAGGLAKLTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MA MEPPEPPDAGATVALVEVREALTH RRRQVQQEARRAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG			1	ł	1		HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE
NMWDLSSPGLINGRIPDESKMAYTHEIRU AMGVVGGVLGAVFMALNVELTMFRIRVIN- AMGVVGGVLGAVFMALNVELTMFRIRVIN- PCLQVIEAVLVAAVTATVAFVLIYSSRDCOJ QGGSMSYPLQLFCADGEYNSMAAAFPHTI- SVVSLFIEDPESVINLTIGLFILTVYFFLACV YGLTVSAGVFPESLLIGAAWGRLFGISLSVI- AAIWADPGKYALMGAAAQLGGIVRMTLSI VIMMEATSNVTYGFPIMLVLMTAKIVOOD GLYDMHIQLGSVPFLHWEAPTVSISLTARE GMSTPVTCLRRREKVGVVDVJLSDTASNHNN PVVEHADDTOPARLQGLLISPULTLEHK: VERSILGLVQRRLRLKDFRDAYPRPPIQSI VSQDERECTMDLSEFANRSPYTVPQEASLT- VFKLFRALGERHLVVVDINROVVGLVTRK LARYRLGKRGLEELSLAQTGFKAQATAEGI AGAAQQPCQLARVTLEDLGLAGLASPPI LSLEELSERYESSHFTSTASVPEQDTAKHVU- LEQWVVELQARVTCLERHKGRCERATRSI. RELLQVRARVQLQGSELRGLQQEARPAAQ, EKEAPFSGLQNOMQALDKRIGGESARTRSI. RELLQVRARVQLQGSELRGLQQEARPAAQ, EKEAPFSGLQNOMQALDKRIVEVREALTI RRRQVQQEAERRGAGDSAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVREALTI RRRQVQGEAERRGAGDSAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVREALTI RRRQVYOEQAFERRGAGDSAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVREALTI RRRQVYOEQAFERRGAGDSAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVREALTI RRRQVYOEQAFERRGAGDSAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVREALTI RRRQVYOEPTDPKFPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVKAAWAQGYTPDFKPPQQWYUS VTQRDLMVAAWAQGYTPDFKPPQQWYUS VTQRDLMVAAWAQGYTPDFKPPQQWYUS VTQRDLMVAAWAQGYTPDFKPPQQWYUS RPAHLNANDWATNGVGRX VSISYYGYQL KPAHLNANDWATNGVGRX VSISYYGYQL KPAHLNANDWATNGVGRX VSISYYGYQL KPAHLNANDWATNGVGRX VSISYYGYQL KPAHLNANDWATNGVGRX VGRAGGLAGLAGLAGLAGLAGLAGLAGLAGLAGLAGLAGLAGL		!					KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG
AMGVVGGVLGAVFNALNYWLIMFRIRYITY PCLQVIEAULVAAVTATVAFUTYSISRDCOJ QGGSMSYPLQLFCADGGYNSMAAAFFNIP PCLQVIEAULVAAVTATVAFUTYSISRDCOJ QGGSMSYPLQLFCADGGYNSMAAAFFNIP SVVSLFIDPPGSVPNLTLIGLFLVYFLACV YGLTVSAGVFIPSLLIGAAWGRLFGISLSYL AAIWADPGKYALMGAAAQLGGIVRMTLY VGLTVSAGVFIPSLLIGAAWGRLFGISLSYL AAIWADPGKYALMGAAAQLGGIVRMTLS UMMEATSNVTYGFIPMLVLMTAKTVGDVI GLYDMHIQLQSVPFLHWEAPVTSHISLTASH MTSTPVTCLRREEKVGVIVDVLSDTASNHNC PVVEHADDTQPARLQGLIRSQLIVLLEWH VERSNLGLVQRRILRLEWFRADAYTRFPIQSI VSQDERECTMDLSEFNMFSPYTVPQEASTAV VERSNLGLVQRRILRLOFPKAQLATAETI LARYPLGREGLEELSLAGTGPKAQATAETI LARYPLGREGLEELSLAGTGPKAQATAETI LARYPLGREGLEELSLAGTGPKAQATAETI LARYPLGREGLEELSLAGTGPKAQATAETI LARYPLGREGLEELSLAGTGPKAQATAETI LARYPLQREVQLQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRGLAGLASLTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRGLAGLTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRGLAGLTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQSELRGLAGLATDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQSELRGLAGLATDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQSELRGLAGLATDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLAGSELRALTDLL QEEQGREVACCALQKNQEDSSRRVDLEVA MELEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLAGSELRALTDLL QEEQGREVACCALQKNGEDSSRRVDLEVA MELEQWVVELQAEVACLAGETACLAGLAGAAVALAQCOKVTINT AVRIPGGPAVANSVARKHGFILNGLEFQLOV MELEQWVENTATRSLAGGETARSLAGGLAGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							1
PCLQVIEAVLVAAVTATVAFVLIYSSRDCQ) QGGSMSYPLQLFCADGEVIRAAAFRIPT SVVSLFHDPPGSYPRLTIGLFTLYYFFLACV YGLTVSAGVFIPSLLIGAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGISLSYI. AAIWADPGKYALMOAAWGRIFGIALWOONG GLYDMHQLQSVPFLHWEAPVTSISLTARE MSTPVTCLRREKVGVIVDVLSDTASNHN PVVEHADDTOPARLQGLIRGUSIVLLKHKV VERSNLGLVQRRLRLKDFRDAYPRFPPIQSI VSQDERECTRULSEFRINGSQLIVLLKHK VERSNLGLVQRRLRLKDFRDAYPRFPIQSI VSQDERECTRULSEFRINGSQLIVLLKHK VERSNLGLVQRLRLKDFRDAYPROVIDLYGLVAL LARYRLGKRGLEELSLAQTGPKAQATAEGI AGAAQQFCQLRAVTLEDLIGLLAGGLASPI LSLEELSERVESSHPTSTASVPEQDTAKHWI LSQWVSLQAVACLERHQRCERATRSL RELLQVRARVQLQGELRQLQQEARPAAQ EKSAPPFSGLONQMQALDKLYGVERATATI RRRQVQQEABRRGAEQBAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854  2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WILWVVAATGTLVLLAADAQGQKYFINT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HEWHRGVTKRSLSPHRPSRLQREPQVQU EQQVAKRRTKRDVYQEPTDFKRPQOWYLS VTQRDLMWKAAWAQGYGFGIGVVSILDD ENHPDLAGNYDPGASFDVNDQDPDPQPR TQMDNHRHGTRCAGEVAANNGVCGVQV AYNARIGGVRMLDGEVTDAVEARSLGLNP HHITYSASWGFEDDGKTVDGPARLAEAPF GVSQGRGGLSIFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGGISFWAKSONGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEGTITARSUNGTRSTLLAARPID SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTILKFILVLYLGTAPFGLEVPPPE GCKTLTTSSQNCVCCEGGSLIQRSCCVQHC GGCTLTTSQQCVCCEGGSLIQRSCCVQHC GGCTLTTSQQCVCCEGGSLIQRSCCVQHC GGCTLTTSQCQCVPPRIPPEYPEYPAGQRLAGGLE SEEDGGCGTTAFIKDQSAL  BSSS 2205 A 7058 3 1441 QRPASQLLAPFAABALPGAFRAAMAQGIFST			ļ				
QGGMSYPLQLECADGEYNSMAARFNTP SVVSLFHPPGSSYNPLIGAAWGRLFGISLSYL AAIWADPGKYALMGAAQLGGIVRMTLSI VIMMEATSINVTGFPIAUVIFLACV GLTVSAGVFPSLLIGAAWGRLFGISLSYL AAIWADPGKYALMGAAQLGGIVRMTLSI VIMMEATSINVTGFPIAUVIMTAKINGDWI GLYDMHQLQSVPFLHWEAPVTSHSILTARE MSTPVTCLRREKKVGVIVDVLSDTASNHNO PVVEHADDTOPARLQGLILSSQLIVLKHK VERSILGLVQRRLRIKDFRDAYPRFPPIQSI VSQDERECTMDLSEFMNPSPYTVPGEASLP VFKLFRALGIERILVVVDNRNQVVGLVTKK LARYRLGKRGLEEISLAQTGFAAQATAEGE AGAAQQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYSSHTSTSTORDOTAKHWI LEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQSERERGAEQEAGLRALKTDLL QEEQGREVACGALQKNIQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAGQGVTINT AVRIPGGPAVANSVARKIGFLNLGQIFGGY HFWHRGVTKRSI.SPHRPRHSRLQREPQVQI EQQVAKRRTKRDVYQEPTDFKFPQQWYLI VTQRDLMVKAAWAQGTGHIVVSILDDI ENHPDLAGNYDPGASFDVNDQDPDQFR TQMNDNRHGTRCAGEVAAVANNGVCGV AYNARIGGVRMLDGEVTDAVEARSIGLNP HHHYSASWGFEDGKTVDFQRALAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYYNSTYLSISSATOFGNVWYSEACSSTIL TTYSSGNQNEKQIVTTDLRQKCTESHTOTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KFAHLMANDWATNGVGRKVSHSYGGLL AGAMVALAQNWTTVAPQRKCITBLITERL AGAMVALAQNWTTVAPQRKCITBLITERL AGAMVALAQNWTTVAPQRKCITBLITERL GKRLEVRKTYTACLGEPHITRLEHAQAK LSYNRRGDLAHLUSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWULEIE TSEANNYGTLTKFTLVLYGTAPGGLPVPFE GCKTLTSSQACVCCEGGFSLHQKSCVQHCF GRAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCFSHASLDFVEGTCSR QSSRSEPPQQPPRLIPPEVAGQRLRAGLI HLPEVVAGLSCAFIVLYFVTVLVLQLRSGF FROWKYTTMMFGGLISYRGTPEAWQEECPE SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAFFAAEALPGAFRAAMAGHFSI ACDVVGFDLDHTLCKYNLPESAPLTYNSSA		[	ĺ	1			
SVYSLFHDPPGSYNPLTIGLFILLYFTLACY YGLTVSAGVPIPSLLIGAGNICHTGISLSYL: AAIWADPGKYALMGAAAQLGGIVRMTLSI VIMMEATSNYTYGFPIMLYLMTAKHYGDVI GLYDMHQLGSVPFLHWEAPYTSHSLTARE MSTPVTCLRREEKVGVIVDVLSDTASNHN PVVEHADDTOPARLQGALIGAUINLENKE WSTPVTCLRREEKVGVIVDVLSDTASNHN PVVEHADDTOPARLQGALIGAUINLENKE VERSNLGLVQRRLRLKDPRDAYPRFPPIQSI VSQDERECTMDLSEFMNPSPYTVPQEASLP, VFKLRRALGLRHLVVVDNRNQVGLVTRK LARYRLGKRGLEELSLAQTGPKAQATAEGI AGAQQPCQLRAVTLEDILLAGGLASPI LSLELSERVESSHPTSTASVPEQDTAKHWI LEQWVSELQAEVACLREHQRCERATRSI RELLQVRARVQLQGSELRQLQQEARPAAQ EKEAPETSGLQNQMQALDKRLVEVREALTI RRRQVQQEARRGASQEARLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPPPDQAKETGAPGSQPPYPMELI WILWVVAATGTLYLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGGYV EQQVAKRRTKRDVVQEPTDFPMELI WILWVVAATGTLYLAADAQGGKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGGYV VOQRDLMVKASWAQGTGHGIVVSILDDL EKNHPDLAGNTYDFASSFDVDDQDPDPQPR TQMMDNRHGTRCAGEVAAVANNGVGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNP HHHYSASWGFEDDGKTVDGPARLAEEAFF GVSQGRGGLGSFVWASGNGGREHDSCNC GYYNSIYTLISSATQFGNYDGPALLAEEAFF GVSQGRGGLGSFVWASGNGGREHDSCNC GYYNSIYTLISSATQFGNYDGHLTEPKL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLLVSNGRGSTLIFERL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLLVSNGRGSTLIFERL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLLVSNGRGSTLIFERL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLLVSNGRGSTLIFERL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLTSNGRGSTLIFERL GKRLEVRKTVTACLCEPHITRLEHAQAKL LSYNRRGDLAHLTSNSMGTSSTLLAARPHD SADGPHDWARMTTHSWGEDPSGGEWLEEL TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEGTCSR QSSRSEPPQQPPRIPPEVEAQGRLRAGLLF HPEVVAGLSCAFIVLYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSYTMORGLISVGLTYFTVFLVLQLRSGF FROXVSTYTMORGLISVGLTYFTVFLVLUSTANDAGHTSI ACDVVGFDLDHTLCXVNLPESAPLTYNSFA		ļ					
### STATES    ##							
AAIWADPGKYALMGAAQLGGIVRMTLSI VIMMEATSNYTYGFINLVLMTAKIVGDVI GLYDMHQLQSVPFLHWEAPVTSHSLTARE MSTPVTCLRREEKGVIVDVLSDTASNHHOK PVVEHADDTOPARLQGLIRSQLIVLKHK VERSINLGLVQRRIRLKDFRDAYPFPPTOSI SQDERECTMDLSEFRHPSPYTVPQEASLP VFKLFRALGIRHLVVVDNRNQVVGLVTRK LARYRLGKRGLEELSLAQTGPKAQATAEGI AGAAQQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREIKQRCERATRSL REILQVRARVQLQAEDRALDREIKQRCERATRSL REILQVRARVQLQAEDRALAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELT RRQVQQEAERRGAEQEAGLRAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELT RRQVQVQEAERRGAEQEAGLRAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELT RRQVQVQEAERRGAEQEAGLRAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  855 2206 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELT RRQVQVQEAERRGAEQEAGLRAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELT RRQVQVQEAERRGAEQEAGLRAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA MHEWRACHATAGATAGATAGATAGATAGATAGATAGATAGATAGATA		}	1	1			YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG
GLYDMHIQLQSVPFLHWEAPVTSHSLTARE MSTPVTCLRREEKYQDVJSDTASNHM PVVEHADDTQPARLQGLILRSQLIVLLKHK VERSNLGLVQRRIRLKDFRDA YPRFPFIQSI SVQDERECTMDLSEFRMPSPYTVPQEASLP VFKLFRALGIRHLVVVDNRNQVGLVTIK LARYRLGKGGLEELSLAQTGPKAQATAEGI AGAAQQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREIKQRCERATRSL RELLQVRARVQLQGSELRQLQQBARPAAQ, EKEAPEFSGLQNOMQALDKRLVEVREALTH RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPVPYPMEL WLLWVVAATGTLVLLAADAQGQKVFINT AVRPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRSBLQREPQVQ) EQQVAKRTKRDVYQEPTDPKFPQQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPQR TQMNDNRHGTRCAGEVAAVANNGVGGV AYNARIGGVRMLDGEYDAVASSAGLGNP HIHIYSASWGPEDDGKTVDGPARLAEAFF GVSQGRGGLGSIFVWASGNGGRUDSCNC GYTNSIYTLSISSATQFGNPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIATLEANKTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQKCIIDLITEPKL GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIFLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAIHLVSPMGRSTLLARPPID SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPGLPVPPE GCXTLTSSQACVVCEGSHAGLDPVEQTCGR QKSRESPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVYLQLRSGI FRGYKVYTMDRGLISYKGLPPEAWGECEP SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAFFAAEALPGAPRAAMQHFISI			}	į			AAIWADPGKYALMGAAAQLGGIVRMTLSLT
MSTPVTCLRREKVGUIVDVLSDTASNHNO PVVEHADDTOPARLQGLIRSQLIVLKHK: VERSNLGLVQRRLRLKDFRDAYPRFPFIQSI VSQDERECTMDLSEFMNPSPYTVPQEASLP VFKLFRALGLRRILVOTNRNQVVGLVTKK LARYRLGKRGLEELSLAQTGPFAQATAEGI AGAAQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREIKRQCERATRSL RELLQVRARVQLQSELRQLQQEARPAAQ, EKEAPETSGLQNQMQALDKRLVEVREALTI RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPRISRLQREPQVQI EQQVAKRRTKRDVYQETDFKFPQQWYLS VTQRDDLMYKAAWQYTGHGIVYSILDD EKNHPDLAGNYDPGASEDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAELAFF GVSQGRGGLGSFVWASGNGGREHDSCNC GYTNSIVTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTVAPQRKCIDLTEPKD GKRLEVRKTVTACLGEPNHTRLEHAQARL LSYNRRGDLAHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLLVLYGTAFEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCCSR QSSRESPPQQQPPRLPPEVEAGQRLRAGLLE HLPEVVAGLSCAFIVLVYVTLVLQLRSGF FGVKVYTMDRGLISYKGLPPEAWQEECP SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASGLLAPFAAEALPGAPRAAMAQHFSI		}	J	J	]	]	VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE
PVVEHADDTOPARLQGLIRSQLIVLIKHK: VERSNLGLVQRRLRLKDFRDAYPRFPPIQSI VSQDERECTMDLSEFMNPSPYTVPQEASLP. VFKLFRALGLRHLVVVDNRNQVVGLVTRK LARYRLGKRGLESLA,OTGPEAQATAEGI AGAAQOPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRQLQCBARPAAQ, EKEAPEFSGLQNMQALDKRLVEVYERALTI RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQFGDY HFWHRGVTKRSLSPHRPRISRLQREPQVOY EQVAKRRTKRDVQFTDVKFPQWVLIS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHFDLAGNYDPGASFDVNDQDPDPQFR TQMNDNRHGTRCAGEVAAVANNOVCGVC AYNARIGGVRMLDGEVTDAVPARSLGLNP HIHTYSASWGPEDDGKTVDGPARLAGEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATOFGRVPWYSEACSSTL TTYSSGNONEKQNTUDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVO; KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIDDLTEPFL GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHILVARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLVGTAPEGLPVPPE GCKTLTSSQACVVCEGEFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCOGPALTDCLSCPHASLDPVEQTCSRC QSSRESPPQQPPRLPPEVBAGRIRAGLLE HLPEVVAGLSCAFIVLVYTVTVLVLQLRSG FRGVKVYTMDRGLISYKGLPPEAWQECCP SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFISI							GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV
VERSNLGLVQRALRI,KDFRDAYPRFPPIQSI VSQDERECTMDLSEFMNPSPYTVPQEASLP VFKLFRALGIRHLVVVDNRNQVVGLVTRK LARYRLGKRGLEELSLAQTGPFAQATAEGI AGAQQPCQLRAVLDEIGILLAGGLASP LSLEELSERYESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREIKRQCERATRSL REILQVRARVQLQGSELRQLQQEARPAAQ EKEAPETSGLQNQMQALDKRLVEVREALTI RRRQVQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M AGTWEPRPYDQAKETGAPGSQPPVPPMELI WILLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGYTKRS1SPRPRISHSLQREPQVQ EQVAKRRTKRDVVQEPTDPKFPQQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNIFPDLAGNYDPGASFDVNDQDPPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGFEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIVTLSISSATJOFNVPWSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLIANDWATNATGYGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIDDLTEPKL GKRLEVKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHLUSPPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEE TSEANNYGTLTKFLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEGGFSLHQKSCVQHCG GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQPALTDCLSCFSHASILDVPVPPES GCKTLTSSQACVVCEGGFSLHQKSCVQHCG GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQPALTDCLSCFSHASILDVPVPPES GSRSSPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTTVLVLQLRSG FGGVKVYTMDRGLSYKGLPPEAWQEECF SEEDEGRGERTAFIKDQSAL			1	!	ì		MSTPVTCLRRREKVGVIVDVLSDTASNHNGF
VSQDERECTMDLSERMNPSPYTVPQEASIP VFKLFRALGLRHLVVVDNRNQVVGLVTRK LARYXLGKRGLEELSLAQTGPKAQATAEGI AGAAQQPCQLRAVTLEDLGLLAGGLASPI LSLEGLSERVESSHPTSTASVPEQDTAKHWI LEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRGLQEARPAAQ, EKAPEPSGLQNQMQALDKRLVEVYREALTI RRRQVQQEAERRGAEQEAGIRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPFVPPMEL WILWVAATGTLVLLAADAQGQKVFTINT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHEPRISRLQREPQVQV EQVAKRITKRDVGPTIDKFPQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTIDAVBASLGLGI HHYSSAWGFEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNONEKQIVTIDLRQKCTESHTGTS. APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGTLTRLEHAQABL LSYNRRGDLAHILVGRGKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIIDILTEPKL GKRLEVRKTVTACLGEPNHITRLEHAQABL LSYNRRGDLAHILVGRGRSVSHSYGYGGL AGAMVALAQNWTTVSTENDVETITRLEHAQABL LSYNRRGGLAHILVGRGRSVSHSYGYGGL GGAPQVLDITHYSTENDVETIRLSVCAPCHA CATCQGPALTDCLSCPSHASLDPVEPDCTCSR QSSRESPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVLVLQLRSG FRGVKVYTMDRGLISYKGLPPEAWGECP SEEDEGRGERTAFIKDQSAL  855 2205 A 7658 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFISI	ļ					Į	
VFKLFRALGERHLVVVDNRNQVVGLVTRK LARYRLGKRGLEELSLAQTGPKAQATAEGI AGAAQQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHFTSTASVFEQDTAKHWI LEQWVVELQAEVACREHKQRCERATRSL RELLQVRARVQLQGSELRQLQQEARPAAQ, EKEAPEFSGLQNQMQALDKRLVEVREALTI RRRQVQQEAFRAGEQAGLAKALTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQQGVFTINT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HEWHRGVTKRSLSPHRPHSRLQREPQVQV EQQVAKRRTKRDVYQEPTDFKPPQQWYLS VTQRDLMVKAJWAQGYTGHGIVVSILDDG EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVFBASLGLEP GYSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSTYTLSSSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFULVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCSPSHASLDPVEOTCSR QSSRESPPQQPPRLPPEVBAGQRLRAGLLF HLPEVVAGLSCAFITLVTVTVFLVLQICRGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGGRGRTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASGLAPFAABALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNIPESAPLITYNSFA		Ì	1	İ			
LARYRLGKRGLEELSLAQTGPKAQATAEGI AGAAQQPCQLRAVTLEDLGLLLAGGLASPI LSLEELSERYESSHFTSTASVPEQDTAKHWI LEQWVVELQAEVACLREHKORCERATRSL RELLQVRARVQLQGSELRQLQQEARPAAQ, EKEAPEFSGLQNOMQALDKRLVEVREALTI RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WILWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFUNLGQIFGDY HFWHRGVTKRSLSPHRPHSRLQREPQVQ) EQQVAKRRTKRDVQEPTDPKPPQWYIS VTQRDLMVKAAWAQGYTGHGIVVSILDDI EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRAGEVAAVANIGVCGVQ AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHHYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFWWASGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLROKCTESHTIGTS. APLAAGGILATLEANKNLTWRDMQHLVVQ KPAHLNANDWATNQVGRKVSHSYGYGLL AGAMVALAQNWTIVAPQRKCIIDILTEPKL GRILEANKNLTWRDMQHLVVQ KPAHLNANDWATNQVGRKVSHSYGYGLL AGAMVALAQNWTIVAPQRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTIVAPGRKCIIDILTEPKL GRILEANKNLTWRDMOPHTGRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCR GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRSSPPQQQPRIPPEVEAGGRLRAGLLP HLPEVVAGLSCAFIVLVFYTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECF SEEDEGRGGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAABALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLTYNSFA							
AGAAQOPCQLRAVILEDĪGLLLĀĞILASPI LSLELSENYESSHYSTASVPEQDTAKHWI LEQWVVELQAEVACLREIKQRCERATRSL RELLQVRARVQLQOSELRQLQOEARPAAQ, EKEAPEFSGLQNQMQALDKRLVEVREALTI RRRQVQQEAERGAEQEGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELF WLLWVVAATGTLVLLADAQQQKVFINT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPHSRLQREPQVQV EQQVAKRRTKRDVYQEPIDPKPPQWYLIS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPOPR TQMNDNRHGTRCAGEVAAVANNGVGOV AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDEKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATOFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTÜLRQKCTESHTGTS, APLAGILALLEANKNLTWRDMQHLVVQ KFAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWITTVAPQRKCIIDILTEPKL GKRLEVRRTVTACLGEPNHTRLEHAQARL LSYNRRGDLAHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFITLVLYGTAPEGLPVPPE GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQOPPRLPPEVEAGORLRAGILF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGI FRGVKVYTMDRGLISYKGLPPEAWQEECP SEEDEGRGERTAFIKDQSAL			}	]	1	j	
LEQWVVELQAEVACLREHKQRCERATRSL RELLQVRARVQLQGSELRQLQQEARPAAQ, EKEAPEFSGLQNQMQALDKRLVEVREALTI RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQQGKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGGY HFWHRGVTKRSLSPHRPRHSRLQREPQVQ\ EQQVAKRRTKRDVVQEPTDPKFPQQWYLS VTQWRDLMVKAAWAQQYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HHHYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS. APLAAGIIALTLEANKNLTWRDMQHLVVQ\ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIDILTEPKD GKRLEVRSTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHLVSPMGTRSTLLARRPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPRLIPPEVEAGQRLRAGLLF HPEVVAGLSCAFIVLVFVTVFLVLQLRSG FRGVKVYTMDRGLISYKGLPPEAWQECCP SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA'							AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP
RELLOVRARVQLOGSELRQLÒQEARRAAQ, EKEAPETSGLQNQMQALDKRIVEVREALTI RRRQVQQEAERRGAEQEAGIRLAKLTULL QEEQGREVACGALQKNQEDSSRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELE WLLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHERPHSRLQREPQVQ) EQQVAKRRTKRDVYQEPTDPKFPQQWYLS VTQWDLMVKAAWAQGYTGHGIVVSILDD EKNHFDLAGNTDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHYSASGGPEDGKTVTDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNTOVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIDDLTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSHLPQKSCVQHCF GFAPQVLOTHYSTENDVETTRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQPPRLIPEVPEAGQRLRAGLLF HLPEVVAGLSCAFFVLVPVTVFLVLQLRSG FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAFFAAEALPGAPRAAMAQHFSI				ļ		-	LSLEELSERYESSHPTSTASVPEQDTAKHWNQ
EKEAPEFSGLQNQMQALDKRÜVEVREALTI RRRQVQQEAERRAGACEAGRLAKLITDLL QEEQGREVACGALQKNQEDSSRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELE WLLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPRHSRLQREPQVQV EQQVAKRRTKRDVYQEPTDPKFPQQWYLL VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNIPDLAGNVDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHITRLEHAQAAL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPRIPPEVEAGRLRAGLLF HLPEVVAGLSCAFFVLVPSTGRURAGLLF HLPEVVAGLSCAFFVLVPSTGRURAGLLF HLPEVVAGLSCAFFVLVPSTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAFAABALPGAPRAAMAQHFSI			i				LEQWVVELQAEVACLREHKQRCERATRSLL
RRRQVQQEAERRGAEQEAGLRLAKLTDLL QEEQGREVACGALQKNQEDSSRRVDLEVA M  854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPRHSRLQREPQVQV EQQVAKRRTKRDVYQEPTDPKFPQQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDODPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKKQIVTTDLRQKCTESHTGTS. APLAAGIALTLEANKNLTWRDMOPILVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLKFTLVLYGTAPEGLPVPPES GCKTLTSRQACVVCEGEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFYTYFLVLQLRSG6 FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGRETAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI	1		l		ŀ	ł	
REST 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELE WLLWVAATGTLVLLAADAQGKVFTNT AVRIPGGPAVANSVARKHGFLNL.GQIFGDY HFWHRGVTKRSLSPHRPRHSRL.QREPQVQY EQQVAKRTTKRDVYQEPTDPKFPQQWYLL VTQNDDNRHGTRCAGEVAAVANNGVCGVG AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHLYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWSGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS. APLAAGIIAITLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRVGKVSHSYGYGLLL AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPHHITRLEHAQARL LSYNRRGDLAHHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPPRLPPEVEAGRLRAGLLF HLPEVVAGLSCAFIVLVFYTYFLVLQLRSGG FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI					1		
854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQGGKVFINT AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPRHSRLOREPQVQV EQQVAKRRTKRDVYQEPTDPKFFPQWYLV VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHTYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVFWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS. APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDLITEPKL GKRLEVRKTVTACLGEPNHITTRLEHAQARL LSYNRRGDLAHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPPRLPPEVEAGQRLRAGILF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI	ļ		ļ	ľ	ļ	ļ	
854 2204 A 7037 139 2604 AGTWEPRPYDQAKETGAPGSQPPVPPMELI WLLWVVAATGTLVLLAADAQGQKVFTNT AVRIPGGPAVANSVARKHGFLNLGJIFGDY HEWHIRGVTKRSLSPHRPRHSRLQREPQVQV EQQVAKRRTKRDVYQEPTDPKFPQQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDD EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC:  GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ' KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKI GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRC QSSRESPPQQPPRLPPEVEAGQRLRAGILF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGETAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA							
AVRIPGGPAVANSVARKHGFLNLGQIFGDY HFWHRGVTKRSLSPHRPRHSRLQREPQVQV EQQVAKRRTKRDVYQEPTDPKFPQQWYLS VTQRDLMVKAAWAQGYTGHGIVVSILDDI EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC. GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTIGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACL.GEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETTRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA	854	2204	Α	7037	139	2604	AGTWEPRPYDOAKETGAPGSOPPVPPMELRP
HFWHRGVTKRSLSPHRPRHSRLQREPQVQV EQQVAKRRTRRDVYQEPTDPKFPQQWYILS VTQNRDLMVKAAWAQGYTGHGIVVSILDDO EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLJKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRC QSSRESPPQQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA							WLLWVVAATGTLVLLAADAQGQKVFTNTW
EQQVAKRRTKRDVYQEPTDPKFPQQWYLX VTQRDLMVKAAWAQGYTGHGIVVSILDDD EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC: GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKLI GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAHHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSR QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFYTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA	ì		ľ		•		AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
VTQVRDLMVKAAWAQGYTGHGIVVSILDDG EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMNDNRHGTRCAGEVAAVANNGVCGVG AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS. APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWITVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GGAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRQ QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA	ļ						HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL
EKNHPDLAGNYDPGASFDVNDQDPDPQPR TQMMDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGINP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC: GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLL AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTITKFTLVYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAR				1	}		
TQMNDNRHGTRCAGEVAAVANNGVCGVC AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF, GVSQGRGGLGSIFVWASGNGGREHDSCNC: GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG							
AYNARIGGVRMLDGEVTDAVEARSLGLNP HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC: GYTNSIYTLSISSATQFGNVPWYSEACSSTIL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRQ QSSRESPPQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFTVTVTVLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG							I
HIHIYSASWGPEDDGKTVDGPARLAEEAFF GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRQ QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGP FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG	1						AYNARIGGVRMLDGEVTDAVEARSLGLNPN
GVSQGRGGLGSIFVWASGNGGREHDSCNC GYTNSIYTLSISSATQFGNVPWYSEACSSTL TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG			1	1		[	HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
TTYSSGNQNEKQIVTTDLRQKCTESHTGTS, APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG		Į	1				GVSQGRGGLGSIFVWASGNGGREHDSCNCD
APLAAGIIALTLEANKNLTWRDMQHLVVQ KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG	1	1	l	l	1	1	GYTNSIYTLSISSATQFGNVPWYSEACSSTLA
KPAHLNANDWATNGVGRKVSHSYGYGLLI AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRC QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAC							TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS
AGAMVALAQNWTTVAPQRKCIIDILTEPKD GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRC QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAC				[	1		
GKRLEVRKTVTACLGEPNHITRLEHAQARL LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRC QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAC					1	1	
LSYNRRGDLAIHLVSPMGTRSTLLAARPHD SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRQ QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG			1		1		
SADGFNDWAFMTTHSWDEDPSGEWVLEIE TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLP HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG				1			
TSEANNYGTLTKFTLVLYGTAPEGLPVPPES GCKTLTSSQACVVCEEGFSLHQKSCVQHCP GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG	} 	1	l	l	ł	Ì	SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
GCKTLTSSQACVVCEEGFSLHQKSCVQHCF GFAPQVLDTHYSTENDVETIRASVCAPCHA CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG				[	1		TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS
CATCQGPALTDCLSCPSHASLDPVEQTCSRG QSSRESPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG				1			GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP
QSSRESPPQQPPRLPPEVEAGQRLRAGLLF HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGF FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG						1	GFAPQVLDTHYSTENDVETIRASVCAPCHAS
HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGH FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG	l				1		CATCQGPALTDCLSCPSHASLDPVEQTCSRQS
FRGVKVYTMDRGLISYKGLPPEAWQEECPS SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG	1			1	1		QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS
SEEDEGRGERTAFIKDQSAL  855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFAG				1	l	Ì	
855 2205 A 7058 3 1441 QRPASQLLAPFAAEALPGAPRAAMAQHFSI ACDVVGFDLDHTLCRYNLPESAPLIYNSFA						,	
ACDVVGFDLDHTLCRYNLPESAPLIYNSFA	855	2205	A	7059	3	1441	
	555	2203	] ^	1038	1	1441	ACDVVGFDLDHTI CRVNI PEGADI TVNGFACE
	l · · · ·	1	1	1	1		LVKEKGYDKELLNVTPEDWDFCCKGLALDL
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		1			1		EDGNFLKLANNGTVLRASHGTKMMTPEVLA
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		<u> </u>	<u>_</u> _		<u>L</u>		EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG
						RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFU DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPFAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGUP WALIFFSFASGTFQLVVLYLFSITTSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS OOOLFPSTLFSWLLE
859	2209	A	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

NO: of N nucl-cotide se	SEQ ID NO: of ceptide seq- nence	Met hod	SEQ ID NO: in	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- pe cotide se seq- ue	eptide seq-	nou		ocknumik	HUCICUUUC	LIZ-ASOMUC ACM PERMITATMA ACM
cotide se	eq-			nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq- u	•		USSN	location		
		i	055N 09/496		corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	ACTICC .		914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
	ĺ	1	914	ng to 111st amino acid		Q=Glutamine, R=Arginine, S=Serine,
ſ	i			residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	l			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ			peptide		/=possible nucleotide deletion, \=possible
	——- <del> </del>			sequence	<u> </u>	nucleotide insertion
l	l	}				HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ
1	1	1				DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP
861 2:	2211	$\overline{\mathbf{A}}$	7161	1220	1003	CPS
801 2.	2211	^	/101	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI
l	l					LTTTRIMFHCTYLFASVCLSLLNTLLSPNCL
862 2	2212	A	7211	665	847	KSAMILQ .
802 2.	2212	A	/211	003	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF
863 2:	2213		7010	004	1072	YKNIQKLSFSNYVYHQNYVFSSDWSYDF
803 2.	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP
	j	ļ				LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA
	l	j				RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ
964	1214	<del>-</del>	7014	046	1610	CIKPNYQSPPKECDYNILANSVA
864 2:	2214	A	7214	845	1619	SDKGGKKADRKNHLRHAFPLLPHRVRERLH
	- 1				,	DPKVPVDADHVQGQDPGRAAHDIHGEDVTE
	l					KVSKDPLAPDEVGDTDEGHDRHGHREVGQR
	ł	ļ				HGHDQEEVAYEERACEGGKFATVEVTDKPV
	1					DEALREAMPKVAKYAGGTNDKGIGMGMTV
	]	]				PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP
l	- 1					PAPSDKSVKIEEREGITVYSMQFGGYAKEAD
į	1					YVAQATRLRAALEGTATYRGDIYFCTGYDPP
865 2:	2215	Ā	7246	559	682	MKPYGRRNEIWLLKT
803 2.	2213	A	7240	228	082	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS
866 2:	2216	$\overline{\mathbf{A}}$	7257	641	1310	LANMAKPRLY TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL
800 . 2.	2216	A	1237	041	1310	
	1	İ				DLKKSDFSTRWQKQRCPVVKSKCRENASPFF
	l					FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCGPCPKNWICYKNNCYQFFDESKN
	(					WYESQASCMSQNASLLKVYSKEDQDLLKLV
	1	- 1	•			KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT
	1	l				IIEMQKGDCALYASSFKGYIENCSTPNTYICM
1.	1					QRTV
867 2	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI
20,		.	,200	131	370	MPFFQTLWLMNANRFCSIFTTTNVANNCWW
ł	ł	l				TPYHCWLSVVVCRCESHGI
868 2	2218	$\overline{\mathbf{A}}$	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP
2		^	1270	3	2,2	KGSCPAGGSRMVSESD*EGRGC*ASYPCAC*
						AGS*WR*GSRPAGRGTPPRSLSHARPP
869 2	2219	A	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR
	/	••			, 222	FLTLCTWLLLLGPGLLATVRAECSQDCATCS
- 1	1	i				YRLVRPADINFLACVMECEGKLPSLKIWETC
	}	!				KELLQLSKPELPQDGTSTLRENSKPEESHLLA
	-					KRYGGFMKRYGGFMKKMDELYPMEPEEEA
	1	1				NGSEILAKRYGGFMKKDAEEDDSLANSSDLL
	ł	ļ				KELLETGDNRERSHHODGSDNEEEVSKRYGG
. 1						FMRGLKRSPQLKEKAKELQKRYGGFMRRVG
1	ł	ł				PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE
	j	i		i	!	SYSKEVPEMEKRYGGFMRF
870 2	2220	A	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG
					-0.0	AGQGRGREGAESGGSRGEGPGSDGRLPATGD
}	ļ	ļ			•	FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT
	l					TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL
ļ	j					KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF
	l					VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD
	l	-				FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF
[	[	ſ				GFIASFMFLLDFITMLYEKRQESQLRKPENTT
	l				'	RAEALTEPLNA
871 2	2221	Ā	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN
.	.	-	-		-	DPGNMSFVKETVDKLLTGFRCFREREAAPRR
	l					ALRGAALPGESEAGDPESLRSSVNADWIQYS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1104	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
		i	• • •	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	· •	/=possible nucleotide deletion, \=possible
		ł		sequence		nucleotide insertion
						DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
•		ł	İ	ŀ	ł	PFIC
872	2222	Α	7413	1061	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC
			l			PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
						CGLIKPCPANYFAFKICSGAANVVGPTMCFED
	1	1	ĺ	[	ĺ	RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ
	ŀ	}				KAFDMYSGDVMHLVKFLKEIPGGALVLVAS
		1	1			YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
	1	1	1	ì		SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE
						GWPELLEMEGCMPPKPF
873	2223	Α	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ
		<u> </u>				DHPGQHCETPSLLKIERKLF
874	2224	A	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISITE WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
	ļ					LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL
		1		]		GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
			}	1	}	AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
		1	1	1		MSSLNLDHWLKGAKREEWEPPPQSPALTHSP
						TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
	[	[	1	}		AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	A	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG
0/5	2223	\ ^	7470	, , , , , , , , , , , , , , , , , , ,	~31	SERHEP*HGGVLFRLGPSAPPGKL
876	2226	A	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
670	2220	''	,,,,,		1	ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	A	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC
			, , , , ,	-		TFKDKVLVAARRNASAVVLYNEERYGNITLP
			1	[	[	MSHAGTGNIVVIMISYPKGREILELVQKGIPV
		1				TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII
						SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
		1			1	GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
						KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL
	ļ					DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
		1				NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
		<u> </u>		<u> </u>	<u></u>	KGDAGENTALLEAGRSDSRHGGPIS
878	2228	Α	7586	315	1232	ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
						RGRMQAACWYVLFLLQPTVYLVTCANLTNG
		j			}	GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
	ŀ					QTFRGKENDTDLDLRYDTPEPYSEQDLWDW
		1			1	LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
	-				1	GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
		ł	ļ			AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK
						TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
		1				
070	2220	+	7605	170	301	YKLVQKVCPDYNYHSDTPYFPSG TESWKLKWWSPTCLDQLNGSAPGNVFIHG
879	2229	A	7605 7612	93	391 659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG
880	2230	A	/012	73	659	GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
						VOETDRILVEKRCWDIALGPLKQIPMNLFIMY
	1	1	1		}	MAGNTISIFPTMMVCMMAWRPIQALMAISAT
	1	1	-			FKMLESSSQKFLQGLVYLIGNLMGLALAVYK
		1	.			COSMGLLPTHASDWLAFIEPPERMEFSGGGL
		1				LL COSMOLLI THASDWLAFILITERWILI SUGGE
	1 000	I <sub>A</sub>	7615	291	1452	SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
991	1 .7.7 7 1		1013	1 271	1732	NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
881	2231	1		1		
881	2231					
881	2231	••				SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL
881	2231					SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
881	2231					SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIIFTICFVPYHAFRFIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST
882	2232	A	7617	67	379	RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
884	2234	A	7638	2640	2861	AVATVGPISVAVGASHVFFQFYKKGKHLSS APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP
885	2235	A	7642	201	455	GLSCHTSHSG PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61 _	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A	7702	242	1298	APSHRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

CCCID	Lerom	1 1/-4	Lero	D-31-4-3	1 10 11 11 11 1	
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	nou	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide					location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ŀ	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ		peptide	1	/=possible nucleotide deletion, \=possible
			ļ	sequence	L	nucleotide insertion
i	!	ł				EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
		1			ļ	GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
1			ľ	Ì		DESTKTKDQILTSRINAVERDLLEPSPADQLG
1		l	ł		ł	NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV
			1			TREPARRLFLFGEEPSKLDQDVLAALECADV
1						DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
1 .			j ·			KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
			1			GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240	A	7711	360	269	
891	2241					RHMPVIPALWEAEVGGLLEPRSSRSAWATE
071	2241	Α	7721	61	1175	KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL
		]		ļ	]	VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
		!				HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN
1		ł				KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
		j				YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
			[			AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
·		Ì				WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
		l	. ,:			KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR
		ľ		::	100	SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
						EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
		l				LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
1		1	1			EAFKEWKEWIIKCLAVLQK
892	2242	$\overline{A}$	7723	2	1650	SAPTAPARPCRAERGSGGGMLALLAASVALA
0,2	2272	1 1	112	2	1000	
Į į		ļ				VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
			]			PAMPMQGGAQSPEEELRAAVLQLRETVVQQ
{						KETLASARAIRELTGKLARCEGLAGGKARGA
1						GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
1						DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ
			l • 1			LRETVVQQKETLASARAIRELTGKLARCEGL
	'					AGGKARGAGATGKDTMGDLPRDPGHVVEQ
						LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
1						FREVLQQRLGELERQLLRKGAELEDEKSLLH
						NETSAHRQKTESTLNALLQRVTELERGNSAF
1 1			ł 1			KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
						ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
]						WGNNPIELLINDKVAQLPLFVSDGKWHHICV
	i					TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
						KPGGVLILGQEQDTVGGRFDATQAFVGELSQ
						FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
1 1						NNVDVFGGASKWPVETCEERLLDL
893	2243	Α	7729	3554	2419	
10,5	2273	Λ.	1127	) <del></del>	4417	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
1						DNYNDTSLVENHLCPATEGPLMASFKAVFVP
1						VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET
; !						FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
					•	LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
1						HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
] {						LFAKVSQGHHNNSLPRCTFSQENQAETHAWF
{			j [	{		TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
}						QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
1			:		1	IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL
1				ł	ł	GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
, 1				ļ		CTGPASLCQLFPSWRRSSLSESENATSLTTF
894	2244	Ā	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
1			50	ا ۲۰۰۰	201	
	İ		<u> </u>			VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
	1				ŀ	SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
1 1	}					VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
1005			77.5			D
895	2245	A	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
						LWLSLFLHAGKEAPHCPRTRPL
896	2246	Α	7754	1	372	SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796		807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID Not of nucleotide peptide sequence (A-Alamine C-Cysteine, Deciminal peptide sequence (G-Alamine C-Cysteine, Deciminal peptide sequence (G-Alamine C-Cysteine, Decided boation) (Part of the Cysteine, Decided sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine C-Cysteine, Decided Sequence) (G-Alamine, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, G-Glyteine, H-Histidine, Graphice, G-Glyteine, H-Histidine, G-Gl	050 10	T CEC TO	1 3 6 - 4	1 650	Predicted	D . 1: 4 1 - 4	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide   seq			L			1	
SSN   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Ocation   Option   Option   Ocation   Opt		1	поц	1 -			
Sequence			1	-	I .	B .	
1							
mainio acid residue of peptide residue of peptide sequence   TaThreonine, V=Vaine, W=Tryptophan, Y=Trystop, x=Vulanow, Y=Stop codon, /=possible nucleotide deletion, V=possible nucleotide d	•	uence	ĺ				
Presidue of   Poptide   Sequence   Y=Tyrosine, X=Ukhown, **Siop codon, / Possible nucleotide (client), *possible nucleotid	uence		1	914			
Poptide							
		1	ł	1		sequence	
903   2253   A   7807   1   384   PPILPWSDGRARESSKCPRSRFPYQVGKMA   VSTYSTSSIM_LAISRISLIPSILSVSTSRRPY   RCDSPTDSQKDMEIPLPPWQERTDESIETER   ARLIVESSKRGMEISPLPPWQERTDESIETER   ARLIVESSKRGMEISPLSPWGDYWATEAKPA   PETENDEWMALRIPAKNKKEQILAPIDLE   PEQUIN YDRLINRESSUDWDIYYWATEAKPA   PETENDEWMALRIPAKNKKEQILAPIDLE   PERQUIN YDRLINRESSUDWDIYYWATEAKPA   PETENDEWMALRIPAKNKKEQILAPIDLE   PERQUIN YDRLINRESSUDWDIYYWATEAKPA   PETENDEWMALRIPAKNKKEQILAPIDLE   PERQUIN YDRLINRESSUDWDIYYWATEAKPA   PETENDEWMALRIPAKNKKEQILAMILLIGK   WRSKCCKGIGKSQILPALVETTRYLDLFTNF   DSENDITRIEFILVEVIGI.SFLENYSFILLEL   WYFSIYLESVALPQL/PINSKTGREATITTHYL   FFIGUYRALYLANWIRRYQTERFYDQIAVVS   GVVQITFCOPFYLYVTKGRSWDDSNADTGL   RSYSSI   STANDVLSPQLKDENSKLRRKLNEVQSFSEA   QTEMYRTLERKLEAKMKEESDYHDLESVVQ   QVEQNLEIMTKRAVKAENHVVKLKQISILL   QAQVSNFQREBALRCQGQSLTINLVA   ELIKSIDNISEVKDEEDS   PETENDESSURGERSGASGREARCQGGRITILAVA   ELIKSIDNISEVKDEEDS   DSPRNRFELIGPITTSTIPTRPGPRPAMEDLDAL   LSDLETTTSHMPRSGAPKERPAPLTPPPSVG   HOPQTIGSGESGSAGSKOCHLYSTVCKPRSPK   PAAPAAPPPSSSSGVLGTGLCELDRILQEINA   TOPNITIDEMSQFPSSKVASGEQKEGSGENKK   RPSIPSSSPSGLPKASATSATLELDRILMASLSJ   PRVONHLPASGYTOPYVSTSNEGSSPSPEPTG   KOSLDTMLGILQSDLSRRGVYTQAKGLCOSC   GREICH   PETENDE   GREICHNIHARGGSLCATCGIPVTGRCVA   GREICHNIH			ŀ	1		1	
VSTYFSTSIMLALSRHSLISPILSVTSRRRPY   RODSPTDSOKOMEDIP-PPWGERTDESIETKR   ARLLYESRKRGMLENCILLSLFAKEHLQHMT   EQUILIYDELINDESIDMOMPTYWATAEAPA   EFENEWMALLRIPEAKNKKKEQRLRAPDLEY   LFEKPR							
RGDSPTDSQKDMEPLPPWQERTDESISTER,   ARLLYPERKRGMLENCILLSLAREH_QHMT   EKQLNLYDRLINEPSNDWDIYYWATEAKPAP	903	2253	A	7807	1	584	
				İ			
		ļ	l	ł	}	l .	
				1			
				1	ł	1	EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
904   2254   A   7813   40   821			1		•	ì	EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
AGARLTGWTMNVPRILGDLSHLLAMILLIGK   IWPSEXCXGIGSQULFAU PRESECTS   IWPSEXCXGIGSSQULFAU PRIVAL   PRIVAL   IWPSEXCXGIGSSQULFAU PRIVAL   PRIVA		ĺ	ĺ	i	ĺ	İ	LFEKPR
AGARLTGWTMNVPRILGDLSHLIAMILLIGK   IWRSKCCKGIGSQU PAL VITYDLDLFTNF   ISIYNTVMKVVPLLCAYVTVYMIYGKFRKTF   DSENDTFRLEFILVPVIGLSFLENYSFILEIL   WTSSYLESVALIPQLIMISKTGAEATITITHYL   FFLGLYRALVLANWIRRYQTENFTYDJAVS   GVVQTIFYCPFYLYYTKGSVESWDDSNADTGL   RSYSSI   LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA   QTEMVRTLERILEAKMIKEESDYNDLESVQ   QVGNLELMTKRAVKAENHVVKLKOEISLL   QAQVSNFQRENEALRCGQGASLTVVKQNAD   VALQNLRVVMNSAQASIEQLVSGAETILNVA   LISKIDRISEVKDEEDS   DSFRINFELIGRFTRIFTRFGFRPAMEDLDAL   LSDLFTTSIRMPRSGAPKERFAEPLTPPPSVG   HOPOTOSGESSGASGNDHLHYSTVCKPRSPK   APAPAPPSSSSGVLGTGLCELDRILQELNA   LSDLFTTSIRMPRSGAPKERFAEPLTPPPSVG   HOPOTOSGESSGASGNDHLHYSTVCKPRSPK   PAPAPAPPSSSSGVLGTGLCELDRILQELNA   TOPNITDEMSGPPSKVASGEQKEDQSSBKK   RPSLFSSPSPGLPASATSATLELDRIAMASI.SD   FRVQNHLPASGPTQPVVSSTNGSPSPPEPTG   KGSLDTMLGLUSSDLSRRGVPTQAKGCSCTALL   GGSSFFKDGAPFCPECVFERFSPRGGFCNQPI   RIKIKMVTALGTMHYBPEHFCVGGCSTALL   GGSSFFKDGAPFCPECVFERFSPRGGFCNQPI   RIKIKMVTALGTMHYBPEHFCVGGCSTALL   GGSSFFKDGAPFCPECVFERFSPRGGFCNQPI   RIKIKMVTALGTMHYBPEHFCVGGCSTALL   GGSSFFKDGAPFCPECVFERFSPRGGFCNQPI   RIKIKMVTALGTMHYBPEHFCVGGCSTALL   GRENFHPDHFTCTFCLRPLTKGSFQERAKPY   CQPCFLKLFG   GRILCENIFHARGSLCATCGLPVTGRCVSA   LGRENFHPDHFTCTFCTCLRPLTKGSFQERAKPY   CQPCFLKLFG   CANADA	904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
IWRSKCCKGISGKSQULPALVFTTRYLDLFINF							
SIYNTVMKVVFILCAYVITVYMIYGKFKKTF   DESENDITRILEFILVPVIGLSE_ENYSFTILEEL     WTFSIYLESVALI_PQLEMISKTGEAETITITYL     FFLGLYRALYLANWIRRYQTENYFQIAVVS     GVVQTIFYCDFFYLYTYKGKSWDDSNADTGI     RSYSSI			1		İ		
DSENDTFRLEFILIPVIGI.SFI.ENYSFILLEIL   WTFSIYLESVALIPQI.FMISKTGAEATITHYL   FFLGLYRALYLANWIRRYQTEAFTITHYL   FFLGLYRALYLANWIRRYQTEAFTYDQIAVVS   GVVQTIFYCDFFYLYVTKGRSWDDSNADTGI.   RSYSSI   SSYSSI   SSYSSI   SSYSSI   CTOMPATILERKLEAKMIKESDSYHDLESVVQ   QVEONLELMTKRAVKAENHTVKLKQEISLL   QAQVSNPQREINEALRCGQGASLTVVKQNAD   VALQNLRVYMNSAQASIEQLVSGAETINLVA   ELIKSIDRISEVKDEEDS   DSPRINFEILG.FRFTRFTFRFGFRFAMEDLDAL   LSDLETTTSHMPRSGAPKERPAEPLTPPSVG   DSPRINFEILG.FRFTRFTFRFGFRFAMEDLDAL   LSDLETTTSHMPRSGAPKERPAEPLTPPSVG   GROUNTEDINSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSGFPSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSGFPSKVASGEQKEDQSEDKK   RFSLPSSPSGULGTGI.CELDRILQELNA   TQPNITDEIMSGFPSSSGULGTGI.CELDRILQELNA   TQPNITDEIMSGFPSPSFETG   RFREGREY-YCRREDI.GLASGLSSPPETGI.CELDRILQELNA   GGSSFFEKDGAFCYCCYCTI.KLFG   GRILCENHHARRGSL.CATCGLPVTGRCVSA   LGRRHIPDPIFTCTTCLRFTLGSSFQERAGKPY   CQCCTLKLFG   GRILCENHHARRGSL.CATCGLPVTGRCVSA   LGRRHIPDPIFTCTTCLRFTLGSSFQERAGKPY   CQCCTLKLFG   TVXNSVLSRSMLSLSSSWDFRHAPITIPET   TYTCKMIEMEQAAGQLAELDLLASMFPGENE   LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI   NMNLDVSDEKMAMFSLACILPFKYPAVLPEI   TVRSVLISSRQQTQLNTDLATLQKHCHGDV   CLINATEWREHASGYVSRDTSSSFTTGSTVQ   SVDLIFTRLWYSSHHIYNKCKRKNILEWAKEL   SLSGFSMGKRQVVCVCGPQSACEEFWARLR   KLNWKRILARIREDPFDGTNDETERQRFSIF   EEKYFSVNOARGNHAMDFGQL YQFLNTKGCG   DVPQMFLW   STSSELLPSKYLFESK   SCKHTJGAPGVVCVCGPQSVCDACLLNPQNQG   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVQGYFGHHCEHRMD   SCKHTJGAPGYTCDCVGGYFGHHCEHRMD		ļ	1			ł	I
WTFSIYLESVALPQLFMISKTGEAETITTHYL   FFLGLYRALYLANWIRRYQTENFYDQLAVS   GVVQTIFYCDFFYLYVTKGSWDDSNADTGL   RSYSS      905   2255   A   7817   1399   881   LSNKDVLSFQLKDENSKLRRKINEVQSFSEA   QTEMVRTLERKLEAKMIKEESDYPLDLESVVQ   QVEONLELMKRAVKAENHVVKLKQEISLL   QAQVSNFQRENEALRCGQGASLTVVKQNAD   VALQNLRVVMNSAQASIEQLVSGAETLINLVA   ELKSIDRISSEVKDEEEDS     906   2256   A   7822   3   1462   DSFRNFELIGRFTRTPTRFGFRFAMEDLDAL   QAQVSNFQRENEALRCGQGASLTVVKQNAD   VALQNLRVVMNSAQASIEQLVSGAETLINLVA   ELKSIDRISSEVKDEEEDS   LSDLETTSIMPRSGAPKERPAEPLTPPPSVG   HQPQTGGESSGASGMXDHLYSTVCKPRSFX   PAAPAAPPFSSSGVLGTGLCELDRLLQELNA   TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK   RFSLFSSFSFGLFKASATSATLELDRLMASLSD   FRVQNNLPASGFTQFPVVALAGRAVBEPFFTG   KGSLDTMLGILQSDLSRRGVPTQAGLCGSC   NKPJAGQVVTALGRA WHPDLETVCRGECGTAL   GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI   RHKMVTALGTHWHPEHFCCVSCCEPFGDEG   FHERGGRFYCRRDFLQJFAPRCQGCQGFILDN   YISALSALWHPDCFVCRECFAPFSGGSFFEHE   GRPLCENHFARRGSLCAGTQFVTGRCVSA   LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY   CQPCFLKLFG   GRPLCENHFARRGSLCAGTQFVTGRCVSA   LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY   CQPCFLKLFG   GRPLCENHFARRGSLCAGTQFVTGRCVSA   LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY   CQPCFLKLFG   GRPLCENHFARRGSLGAGTQLAGTARTHYPPET   YTCPKMERQAGAGQLAGLDLLASMFPGENE   LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI   NMNLDVSDEKMAMFSLACILPFKYRAVLPEI   TVRSVLSSRSQQTQLNTDAFLQKHCHGDV   CILNATEWYREHASGFVSRDTSSSFTGSTVQ   SVDLFFTELWYSHHTYNCKKRNILEWAKEL   SLSGFSMPGKPQVVCVCGPQSACEEFWARLR   KLNWKRILRHREDIPFDGTNDETERQRKFIF   EEKYFSVNOARGNHMDFGQLYQLFLNTKGCG   DVPQMFLW   SVDLFFTELWYSHHTYNCKKRNILEWAKEL   SLSGFSMPGKPQVVCVCGPQSACEEFWARLR   KLNWKRILRHREDIPFDGTNDETERQRKFIF   EEKYFSVNOARGNHMDFGQLYQLFLNTKGCG   DVPQMFLW   STSRNYAGRGNHMDFGQLYQLFLNTKGCG   DVPQMFLW   STSRNYAGRGNHMDFGQLYQLFLNTKGCG   DVPQMFLW   STSRNYAGRGNHMDFGQLYQLFLNTKGCG   DVPQMFLW   STSRNYAGRGNHMDFGQLYQLFLNTKGCG   SCRHLJRGAFGCVVTNACASGPCPPHANCRDL   WQTFSCTCQPGYYGPGCVDACLLNPQNQG   SCRHLJRGAFHGYTCDCVQGYFGHHCEHRMD   SCRHLJRGAFGCVVTNACASGPCPPHANCRDL   WQTFSCTCQPGYFGHHCEHRMD   SCRHLJRGAFGCVVTNACASGPCPPHANCRDL   WGTFSCTCQPGYFGHHCEHRMD   SCRHLJRGAF		ĺ	1	i			
PFLGLYRALYLANWIRRYQTENFYDQIAVVS	•	1	l		[		
GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL   RSYSS1		1	l	į.	1	İ	
905   2255   A 7817   1399   881   LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA QTEMVRILERKLLAKMIKEESDYHDLESVVQ QVEQNLELMTKRAVKAENITVVLKQEISLL QAQVSINFQREMEALIRAGQGASLITVVKQNAD VALQNLRVVMNSAQASIEQLVSGAETLNLVA EILKSIDRISEVKDEEEDS     906   2256   A 7822   3   1462   DSPRNRFELIGRPTRTPTRFGPRPAMEDLDAL LSDLETTTSHMPRSGAPKERPAEPLTPPPSY QHQPTGSESSGASGDHLYSTVCKPRSPK PAAPAAPPPSSSSGVLGTGLELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSSPPETEG KGSLDTMIGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGCCSTAL GGSSFFEKDGAPFCPCCYFERFSPRCGFCNQPI RHKWYTALGTHWHPEHFCCVSGEPFGDEG FHEREGRPYCRRDFLQLFAPRCGGCQGPILDN YISALSAL WHPDCFVCRECPAPPSGGSPFEHE GRPLCENHFHARRGSLCATGGLPVTGRCVSA LGRRFHPDHFTCTFCLRFLTKGSFQERAGKPY CQPCFLKLRG     907   2257   A 7828   1792   1671   FIVVNQSFAPSPQEVGTLYECFGSDGKLVLH YCKSQAWG     908   2258   A 7842   110   1172   KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCFKMEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELDCIEKKTMEGRSSKYYFTI NMLDVSDEKMAMPSLACILPPKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWYREHASGYVSRGHSSPTGTSTVQ SVDLIFTRLWTSHHIYNKCKKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFFGTNDETERGRFSISFTGSTVQ SVDLIFTRLWTSHHIYNKCKKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFFGTNDETERGRFSISF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV     909   2259   A 7870   3067   2923   EGICVYTFIYVMTTRTCIMTTYPYMYMNSV LISSELLPSKYLFESK     910   2260   A 7884   212   4874   GALTWSHILAACCPQVWIGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGFGCVDACLLNPCQNGG SCRRLIPGAPHGYTCDCVGGYFGHHCEHRMD		ł	1	i	ļ	l	
905   2255   A   7817   1399   881	•	ļ		ł			
			<u> </u>	L			
0   0   0   0   0   0   0   0   0   0	905	2255	A	7817	1399	881	
						ł	
906 2256 A 7822 3 1462 DSPRNRFEILGRPTRPTRPGPRPAMEDLDAL LSDLETTISHMPRSGAPKERPAEPLTPPSYG HOPOTOSGESSGASGASGMANDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEOKEOQSEDKK RPSLPSSPSPGLPKASATSATLELDRIMASLSD FRVQNHLPASGPTQPFVVSSTNEGSPSPPEPTG KGSLDTMLGLQELNA TQFNITDEIMSQFPSSKVASGEOKEOQSEDKK RPSLPSSPSPGLPKASATSATLELDRIMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLQELDRA SGRIGAR WHPEHFVCGGCSTALL GGSSFFEKDGAPFCPECVFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSGCGSFTGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECGPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECGPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGSFFEHE GRPLCENHFHARRGSLCATGGLPVTGRCVSA LGRRPHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYYNQSFAPSPPQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVYNDQLAVAELKGEKKTMEGRSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASQVSRDTSSSPTTGSTVQ SVDLIFTRL WTYSHHYNKCKRKNILEWAKEL SLSGPSMPGKPQVVCVEGPQSACEEFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVVGARGMHMDFGQLYQFLNTKGCG DVFQMFLWW SKYLLFRIKEDIPFDGTNDETERQRKFSIF EEKVFSVVGARGMHMDFGQLYQFLNTKGCG DVFQMFLW WTYSHTYSHTYNKGCG DVFQMFLW WTYSHTYNKGCG DVFQMFLW WTYSHTYNKGCG DVFQMFLW WTYSHTYNKGCG DVFQMFLW WTYSHTYNKGCGD SCRELPGAPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVAGLLNPCQNQG SCRELPGAPCHCEHNMD			}	1	}	j	
BILKSIDRISEVKDEEDS		•		1			
906			1	l			VALQNLRVVMNSAQASIEQLVSGAETLNLVA
LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPSSSGVLGTGLCELDRLLQELNA TQPNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSSPPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPJAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYSERSPRGGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHERGRPYCRRDFLQLFAPRCQGCQPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHFET YTCPKMEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQTQLINTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRL WTSHHIYNKCKKKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVPQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 1LISSELLIPSKYLFFSK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTFSGSFALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTPSCTCQPGYYGGCVDACLINPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							EILKSIDRISEVKDEEEDS
HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA TQPNITDEIMSOFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPFVVSSTINESSPSPEPPTG KGSLDTMLGLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RIKKMYTALGTHWHPEHFVCCYSCEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNOSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSFTTGSTVQ SVDLIFTRLWTYSHHYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLINWRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYTLFESK  910 2260 A 7884 212 4874 GALTWSHFLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	906	2256	A	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
PAÄAPPPSSSQU-GTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSSPSGLFKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRGGVPTQAKGLCGSC NKPIAGQVTALGRA WHPPEHFVCGGSCTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENIFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTOLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNOARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRKVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD				1			LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
PAÄAPPPSSSQU-GTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSSPSGLFKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRGGVPTQAKGLCGSC NKPIAGQVTALGRA WHPPEHFVCGGSCTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFVCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENIFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTOLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNOARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRKVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD				İ			HOPOTGSGESSGASGDKDHLYSTVCKPRSPK
TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVVSSTNEGSPSPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPJAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKGPYCRROFIQLFAPRCQGCQGPIDN YISALSALWHPDCFVCRECFAPPSGGFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLIP 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPSGLKQSSHLSLSSSWDFRHAPTHFET YTCYKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKOLIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGTVSRDTSSSPTIGSTVQ SVDLIFTRLSTHYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIFFDGTNDETERQRKFSIF EEKYFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNISV LISSEILLPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRKNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRILPGAPHGYTCDCVGGYFGHHCEHRMD			•	1	1		
RPSLPSSPSQLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPEYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVRFHASGYVSRDTSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1		1		
FRVQNHLPASGPTQPPVVSSTNEGSPSPPPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIVYNGSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYFAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWTYSHHYNKCKRKNILEWAKEL SLSGFSMPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			l	1	ì	l	
KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC   NKPIAGQVVTALGRA WHPEHFVCGGCSTAL   GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI   RHKMVTALGTHWHPEHFCCVSCGEPFGDEG   FHEREGRPYCRROFLQLFAFRCQGCQGFILDN   YISALSALWHPDCFVCRECFAPFSGGSFFEHE   GRPLCENHFHARRGSLCATCGLPVTGRCVSA   LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY   CQPCFLKLFG     907   2257   A   7828   1792   1671   FTYVNQSFAPSPDQEVGTLYECFGSDGKLVLH   YCKSQAWG     908   2258   A   7842   110   1172   KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG   SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET   YTCPKMEMEQAEAQLAELDLLASMFPGENE   LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI   NMNLDVSDEKMAMFSLACILPFKYPAVLPEI   TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV   CILNATEWVREHASGYVSRDTSSSPTTGSTVQ   SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL   SLSGFSMPGKPGVVCVEGPQSACEEFWARLR   KLNWKRILIRIREDIPPDGTNDETERQRKPSIF   EEKVFSVNGARGNHMDFGQLYQFLNTKGCG   DVFQMFLWV     909   2259   A   7870   3067   2923   EGICVYTFIYVHMYTRTCMHTYPYMYMNSV   LISSELLIPSKYLFESK     910   2260   A   7884   212   4874   GALTWSHPLLAVCPQGVWLGSTPSGSPALLP   PSHRVNAEPGCVVTNACASGPCPPHANCRDL   WQTFSCTCQPGYYGPGCVDACLLNPCQNQG   SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD					ļ		
NKPIAGQVVTALGRA WHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCGGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRRNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			l				
GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSPTTGSTVQ SVDLIFTRLWYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCADL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			i	1			
RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGFILDN YISALSAL WHPDCFVCRECFAPFSGGSFFEHE GRPLCENIFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIVYNGLFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSFTTGSTVQ SVDLIFTRLWIYSHHIYNKCKKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGCVYTTIYVHMYTRTCMHTTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1	ſ	1		1	
FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRPHPDHFTCTTCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILFFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWTYSHHYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIFFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD				i		1	1
907 2257 A 7828 1792 1671 FIVNQSFAPSPDQEVGTLYECFGSDGKLVLH 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP 911 PSHRVNAEPGCVVTNACASGPCPPHANCRDL 912 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP 913 PSHRVNAEPGCVVTNACASGPCPPHANCRDL 909 2259 WYTSCTCQPGYYGPGCVDACLLNPCQNG 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP 911 PSHRVNAEPGCVVTNACASGPCPPHANCRDL 912 WYTFSCTCQPGYYGPGCVDACLLNPCQNGG 913 SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1	1				
GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG  907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		ļ	)	j	j		,
LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG				i			1
907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		ì					1
907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG  908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		ł	1	1			
908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD				1	<u> </u>		1 -
908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	907	2257	Α	7828	1792	1671	
SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD						L	1
SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	908	2258	Α .	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWTYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1			1	
LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWTYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			ĺ	i	1	1	
NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD					]	1	
TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1		1			
CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1	1			
SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			[	1	ł	1	
SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV  909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK  910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	l		l	1	1	1	
KLNWKRILIRHREDIPFDGTNDETERQRKFSIF				1		1	
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		}	1		l	1	
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1		1	1	1	
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1	1	l		, ,
910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			<b> </b>		1		
910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	909	2259	A	7870	3067	2923	
PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		L	<u></u>	<u> </u>	<u></u>		
WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	910	2260	A	7884	212	4874	
SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	ı		ĺ	1	1	1	PSHRVNAEPGCVVTNACASGPCPPHANCRDL
SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1	l	1		1	
					i	1	
		]	J	1	j	1	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Ì				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			•	residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
<del></del>				Sequence		TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST
						SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV
						TASGCRVLYDACPKSLRSGVWWPQTKFGVL
1						ATVPCPRGALGLRGAGAAVRLCDEAQGWLE
						PDLFNCTSPAFRELSLLLDGLELNKTALDTME
						AKKLAQRLREVTGHTDHYFSQDVRVTARLL
					•	AHLLAFESHQQGFGLTATQDAHFNENLLWA
						GSALLAPETGDLWAALGQRAPGGSPGSAGLV
						RHLEEYAATLARNMELTYLNPMGLVTPNIML   SIDRMEHPSSPRGARRYPRYHSNLFRGODAW
}					'	DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS
						VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ
						AERRGARLPQNPVMNSPVVSVAVFHGRNFLR
						GILESPISLEFRLLQTANRSKAICVQWDPPGLA
						EQHGVWTARDCELVHRNGSHARCRCSRTGT
						FGVLMDASPRERLEGDLELLAVFTHVVVAVS
						VAALVLTAAILLSLRSLKSNVRGIHANVAAA
	•					LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR
					,	FYHALGWGVPAVLLGLAVGLDPEGYGNPDF
						CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA
						ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA
	<u> </u>		· ·			SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV
						LLLFCVLNADARAAWMPACLGRKAAPEEAR
						PAPGLGPGAYNNTALFEESGLIRITLGASTVSS
				1		VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS
						AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF
				·		QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA
						LGECEAAPCALQTWGSERRLGLDTSKDAAN
					ā .	NNQPDPALTSGDETSLGRAQRQRKGILKNRL
ļ						QYPLVPQTRGAPELSWCRAATLGHRAVPAAS
j						YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR
						QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG
						RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR
ļ						DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP
	]					LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ     VPSRHPSREALGPLPQLLRAREDSVSGPSHGP
					!	STEQLDILSSILASFNSSALSSVQSSSTPLGPHT
						TATPSATASVLGPSTPRSATSHSISELSPDSEPR
	1					DTQALLSATQAMDLRRRDYHMERPLLNQEH
1		1		İ	l	LEELGRWGSAPRTHQWRTWLQCSRARAYAL
	1	1		[		LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA
	1			1		IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY
911	2261	Ā	7890	21	806	FLFGTSRHISVESLCVPGPVDT  EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG
711	2201	^	/070	41	800	SCRPKARSSSARWALTCCLVLLPFLAGLTTYL
	1					LVSQLRAQGEACVQFQALKGQEFAPSHQQV
		]				YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP
				[		ALHWEHELGLAFTKNRMNYTNKFLLIPESGD
	1	[		[		YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV
						VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ
		١				PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE
-		<u> </u>				DKTFFGAFLL
912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL
						LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ
1						PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL
	1	1				YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP
1	1	ł		l		ELELALFLVQEPHVWGQTTPKPGKMFVLRSV
	<u> </u>	<u> </u>	L	<u> </u>	L	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHKWIIAPKGFMANYCHGE
913	2263	A	7892	15	849	CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD ECGCG ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL
				·		YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN RLSNITRPFFSKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	A	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK INCSWFIRANPGEITTISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		<b> </b>	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	1	}	peptide		/=possible nucleotide deletion, \=possible
ļ	<del></del>	<b></b>		sequence		nucleotide insertion
	]	]				VILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDSHAPLTVVSSSGOIRVHFCADKV
		l	-			NAARGFNATYQVDGFCLPWEIPCGGNWGCY
				,		TEQQRCDGYWHCPNGRDETNCTMCQKEEFP
	•	Ì				CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
	ł	l			ļ	CQPGNFHCKNNRCVFESWVCDSQDDCGDGS
						DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
						CTCKLYSLRMFERRSFETQLSRVEAELLRREA
	i	İ				PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
						RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
	j	ļ				RSRHSGSLALVSADGDEVVPSQSTSREPERNH
	}	1				THRSLFSVESDDTDTENERRDMAGASGGVAA
		İ				PLPQKVPPTTAVEATVGACASSSTQSTRGGH
						ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ
		Į			ŀ	GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG
						REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
						ASDQGQGLRQPYNATNPGVRPSNRDGPCERC GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
'''	===0	l	///	10/4	1035	GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQQTEDVLAA
		]	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	.,	0.2	GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
						VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE
,			}		i	HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
		ļ			ļ	RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS
						FSSSWPSAQHLTPSVFNPW
921	2271	Α	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF
		ŀ				MMVYSFRALSFKESTWATFQHGGEATKSRSL
000	0000	<u> </u>			1.50	SSTQ
922	2272	Α	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
						VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN WELVKPN
923	2273	A	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
723	22/3	^	7501	1	3023	LLPLLLLPAGCRALEETLMDTKWVTSELAWT
	]	ļ	] .			SHPESGWEEVSGYDEAMNPIRTYQVCNVRES
						SQNNWLRTGFIWRRDVQRVYVELKFTVRDC
						NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
	ļ					MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
		1	}			FGPLSKAGFYLAFQDQGACMSLISVRAFYKK
						CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
		1	]			AVEVSVPLKLYCNGDGEWMVPVGACTCATG
		1				HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
		}	}			NSRTTSPAASICTCHNNFYRADSDSADSACTT
						VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
						LLYNVICKKCHGAGGASACSRCDDNVEFVPR
		1				QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
	1					GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
	] .	j	<b>j</b> .			IASTVTSQMNSVQLDGLRPDARYVVQVRART
						VAGYGQYSRPAEFETTSERGSGAQQLQEQLP
,						LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS
		1				DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA
	I	<b>!</b>				VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
1	ſ :			i	i	TODOD DELETE A TOTAL DELETE OF THE TOTAL DELETE OF TOTAL DELETE OF THE TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELETE OF TOTAL DELET
1	•					KOPGKKEVFVAIKTLKVGYTERORRDFLSFA
						KQPGRREVFVAIKTLKVGYTERQRRDFLSEA SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
						SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
						SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM KYLSEMNYVHRDLAARNILVNSNLVCKVSDF
					·	SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM

nucleotide sequence   USSN   Ocasion   Corresponding   Sequence   USSN   Ocasion   Corresponding   Sequence	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1							
Peptide   Pept	4 -				ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
peptide				ļ	1		
	ļ	İ	]			sequence	
Section   Sect							nucleotide insertion
SLKVIASAGSGMSQPLLDRTVPDYTTETTUG   WLDADKMGRYKESFVSAGFASPDLVAQMTA     EDLLRIGVTLAGHQKKLLSSIQDMRLQMNQT     LPVQV							
			1				SI KVIASAOSGMSOPI I DRTVPDVTTFTTVGD
LPVQV					ļ		
924   2274   A   7985   1   503		1			] -	•	, , , , , , , , , , , , , , , , , , , ,
QLMRELORTEDKKAEDILAAEVISTVITLS   PORVERI,QKIQNAYSKCEYSDDKVQLAM   QTYEMYDKHIRRI.DADIA.RFEADLKDKMEG   SPEESSGGGLKKGRGQKEKRGSRGGRRTS   EEDTPKKKKHKGG   SPEESSGGGLKKGRGQKEKRGSRGGRRTS   EEDTPKKKKHKGG   LPCSPCAQCMSSTERVULQQSHFHNPRWSR   SPIRCYQGHWPHCYHC   SPIRCYQGHWPHCYHC   GPCKVCCITLAIMI.QCHSFTRKDVQVVEHPKS   LNPKYSQEPSI-SADMALKRKCLISISDLDFW   IVDAQPVGIMQTILQNLKKPPPGGFWSQAFQI   RDTQPILPI.GGRYTITRQ   RDTQPILPI.GGRYTITRQ   RDTQPILPI.GGRYTITRQ   SPIRCYQGHWPHSSADMALKRKCLISISDLDFW   IVDAQPVGIMQTILQNLKKPPPGGFWSQAFQI   RDTQPILPI.GGRYTITRQ   RDTQPILPI.GGRYTITRQ   SPIRCE   SPIRCE   SAMALKRCLISTSDLDFW   IVDAQPVGIMQTILQNLKKPPPGGFWSQAFQI   RDTQPILPI.GGRYTITRQ   SPIRCE   SPIRCE   SAMALKRCLISTSDLDFW   IVDAQPVGIMQTILQNLKKPPPGGFWSQAFQI   RDTQPILPI.GGRYTITRALSV   SPIRCE   SPIRCE   SPIRCE   SAMALKGAR   SPIRCE   SPIRCE   SPIRCE   SAMALKGAR   SPIRCE   SP	024	2274		7085	1	503	
PPORVERLOKIQNAYSKCKEYSDDK/QLAM	924	22/4	A	/963	1	505	
SDFESSGGRILKERGQKEKRGSRGRGRTS     SPECTPTKKKKKHGG			ļ				PDQRVERLQKIQNAYSKCKEYSDDKVQLAM
P25   2275   A   7994   447   589	ļ		l	ľ			
925   2275   A   7994   447   589	i						
926	925	2275	A	7994	447	589	
LINPKYSQIENFISADMALKRKCLISISDLDFW   IWDAQPVGIMQTLQNIKKIPNGGCFWSQAFQI   RDTQPILPLGGRYYTITRQ     927   2277   A   7998   2   353   RIQRPLNSRSPNHSLFVKAELTAQATMKISV     CLLIVTLALCCYQANAFCPALVSELLDFFFI   SEPLFKLSLAKFDAPPEAVAKLGVKRCTDQ   MSLQKRSLIAEVLVKILKKCSV     928   2278   A   8004   130   588   LAPLRCQFGTRTQPRSHPAANDPSAAMSAAG   ARGLRATYHRLIDKVELMLPEKLRPLYNHPA   GPRTVFFWAPMK WCLAGLADMARPAEK   LSTAQSAVLMATGFIWSRYSLVIPKNWSLFA   VNFFVQAAGASQLFRIWRYNQELKAKAHK     929   2279   A   8007   2   1016   EFARRRVFIAAREMSLTSLRVFLVARTGSYP   AGSLLRQSPQPRTITYTAGFRLSAASSKELLM   KLRKTGVFFVORKALETCGGDLKQAEWL   HKEAQKEGWSKAAKLQGRTKEGLIGLLQE   GNTTVLVEVNCETDFVSRNILKFQLLVQVAL   GTMMHQQTILKDQPSAYSKGFLNSSELSGLPA   GPDREGSLKDQLALAIGKLGENMLKRAAWV   KVPSGFYVGSVVHQAMQSPSLHKLVLQKYG   ALVICETSEQKTNLEDVGRRLGQHVVGMAPL   SVGSLDDEPGGEAEKTMLSQPYLLDPSITLGQ   YQPQQYSVVDFVRFECGGEGEAAETE     930   2280   A   8008   3   1679   NSRVWGPWTEPSAGSLRPMARKQNRNSKEJ   SVGSLDDEPGGGRAFKMLSQPYLLDPSITLGQ   YQPQQYSVVDFVRFECGGEGEAAETE   TLLWSVTVSIFAIGGL VGTLIVKMIGKVLGRK   HTILLANNGFAISSAALLMACSLQAGAFFMLIV   GRFMGGGVALSVHYLSEFKREGSLG   QVTAIFICIGGVTGQLLGLECHLRSGVWPLF   GVTVVPAVVQLLSPFLDSPRYLLLEKHNEA   RAVKAFQTFLGKAGTPSPRYLSEFKRERGSLG   QVTAIFICIGGFTGQLLGLECHLEKESTWPYLF   GVTVVPAVVQLLSPFLDSPRYLLLEKHNEA   RAVKAFQTFLGKAGTPSPRYLTSTGG   IETLAAVFSGLVEHLGREPLLIGGFGLMGLFF   GTLTITLILQDHAPWYPYLSEISKERGSLG   GTLTITLILQDHAPWYPYLSEISKERGSLG   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   GTLTITLILQDHAPWYPYLSTGGFGLMGLFF   G		L	<u> </u>		L		
1	926	2276	A	7996	925	582	
RDTQPILPLGGRYYTTIRQ					}		
CLLLVTLALCCYQANAEFCPALVŠELLDFFF    SEPLFKLSLAKFDAPPEAVAKKLGVKRCTDQ    MSLQKRSLIAEVLYKLILKKCSV    MSLQKRSLIAEVLYKLILKKCSV    A	]	1			İ		
SEPLEKLSLAKFDÅPPEAVAAKLGVKRCTDQ   MSLQKRSLIAEVLVKILKKCSV     928   2278   A   8004   130   588	927	2277	Α	7998	2	353	
MSLQKRSLIAEVLVKILKKCSV				1			
928   2278   A   8004   130   588					1	1	
GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVGAAGSQLFRIWRYNQELKAKAHK  929 2279 A 8007 2 1016 EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHITYAGPRLSASASSKELLM KLRRKTGYSFYNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETIDFYSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEGKTNLEDVGRRLGOHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTFPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFINGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPSSRYLLLEKHNEA RAVKAFQTEKKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIEGKAGPPAKIPYVTLSTGG IETLAAVFSGLVEHL.GRPELJLGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFFFIGGSLG YYPLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKNGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	928	2278	A	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG
LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA							-
929 2279 A 8007 2 1016 EFARRRVTIAAREMSLIRSILVFLVARTGSYP AGSLLRQSFQPRHTFYAGPRISASASSKELLM KLRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGKKTKEGLIGLLQE GNTYVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEFGGEAETKMLSQPYYLDPSITLGQ YVQPQGVSVVDFVRFECGGEEAAETE  930 2280 A 8008 3 1679 NSRVWGFWTEFSAGSLRFMARKQNRNSKEL GLVPLTDDTSHAGPPGFGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTYNKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFFMGIDGGVALSVLFMYLSEISFKEIRGSLG QVTAFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVEHLIGRRPLILGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFTQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICTTGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRQKFGYSVNRKRL NNAARKAAPRIECSHIRHAWDHAKSVRQNL				ļ			
AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSVVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFFECGEGEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGHNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITITLIQDHAPWVPYLSIVGLAILASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		}					
KLRRKTÖYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GGIVVYAVVQLLSLPFLPDSPRYLLLEKHNEA RAYKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPLLIGGFGLMGLFF GTLITTLTLQDHAPWVPYLSIVGILAITASFCSG PGGIPFILTGEFFQQSQRPAAFIAGTVNWLSN FAVGLLFPFFQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	929	2279	Α	8007	2	1016	I .
HKEAQKEGWSKAAKLQGRKTKEGLIĞLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAFKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFFECGEGEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFTGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAILASFCCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKNIGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			į			•	
GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNILEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRTIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			į				
GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFEMLIV GRFFMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGREPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1				1		
KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE  930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAHFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					İ	ł	
ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGYFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPFLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			}		}		
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGFGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	]	ļ	ļ	ļ	1	ļ	ALVICETSEQKTNLEDVGRRLGQHVVGMAPL
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
GLVPLTDDTSHAGPPGGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	030	2280		8008	3	1679	
LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	330	2200	1	0,000		10.7	
TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL, YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1		ŀ		1		
HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL, YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			1	1			
GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGGRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		ļ	ł	ł	l	1	
RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					1		
IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL				Į			
CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		-		ļ			IRLVSVLELLRAPYVRWOVVTVIVTMACYOL
GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					1		CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG
PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			]	İ			
FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL , YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL						1	PGGIPFILTGEFFOOSORPAAFIIAGTVNWI.SN
YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP  931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1					}	
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL						ļ	YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI
NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	031	2201		8000	961	300	
	731	2201		0003	001	300	
<u></u>		<u> L</u>	<u>l</u>	<u></u>		<u> </u>	

000 10	CEO ID	N/a+	CEO	Deadistad	Dradioted and	L Amino coid consume (A-Al-i C-C
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted   beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		i		peptide	ł	/=possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
		İ				PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
		j				LIDYVRYMVENHGEDYKAMARDEKNYYQD
						TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
020	2202		0011	410	,	MEVE
932	2282	Α	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
		ĺ	l			DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
						QGRGQIPIPCPWPPPPPPPPPPGSPGPGCRQFHQ SLEAKARHPASVREMRGKVKMRRALRRAPA
•			]			STRASSROPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
755	2203	\ ^`	0012	<b>^ * * * * * * * * * *</b>	10,,	ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
					1	NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
		1			1	RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
		1	[			PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
			}	İ	1	NASQLITQRAQVSLLIRRELTERAKDFSLILDD
	1		1			VAITELSFSREYTAAVEAKQVAQQEAQRAQF
				Ì		LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
	}	İ		ł	1	SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
						DNLVLNLQDESFTRGSDSLIKGKK
934	2284	Α	8023	255 .	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
				]		RLRKFRELHLMRNEARKLNHQEVVEEDKRL
		1				KLPANWEAKKARLEWELKEEEKKKECAARG
	}					EDYEKVKLLEISAEDAERWERKKKRKNPDLG
				ŀ		FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
		Ì				KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE KQIEKRDKYSRRRPYNDDADIDYINERNAKF
		]		1		NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
,,,,		}	552.	1		QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
					ļ	SQHSSPAPMYSQTFHILVLG
936	2286	A	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
				1		FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
	<u> </u>	Ì				IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
						NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
					1	DVNDKRQVSKERGEKLALDYGIKFMETSAK
	ĺ	Ì	İ			ANINVENAFFTLARDIKAKMDKKLEGNSPQG
027		<del></del>	0020	202		SNQGVKITPDQQKRSSFFRCVLL
937	2287_	A	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	A	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
		J	}		)	PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
						ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
					ł	LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
						GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
	1		[	]		VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	A	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
		1			1	AEQLKWSAELARLGESIMDGKQGGMDGSKP
		]	Į.	1	1	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
1	]			1	·	IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
	1		1	1		CLGGHLSCVKILLKHGAQVNGVTADWHTPL
	1	1		1	Į	FNACVSGSWDCVNLLLQHGASVQPESDLASP
	ĺ	1		1		IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
	ŀ	1		1	i	PLYLACENQQRACVKKLLESGADVNQGKGQ
	l	1		1	1	DSPLHAVARTASEELACLLMDFGADTQAKN
	1	1		1	1	AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
	1	1				CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
					i .	
-						L
940	2290	A	8058	2	1203	L KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
·		-				ANSVVVWVNIQAKTTGYDTHCYILNLAIADL WVVLTIPVWVVSLVQHNQWPMGELTCKVTH LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS RKKMVRRVVCILVWLLAFCVSLPDTYYLKT VTSASNNETYCRSFYPEHSIKEWLIGMELVSV VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI PFTCRLEHALFTALHVTQCLSLVHCCVNPVL YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA SRVSETEYSALEQSTK
941	2291	Α	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLIATANQGSEATYESGF VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	Α	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF
944	2294	A	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM
945	2295	A	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE
946	2296	Α	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
948	2298	В	8093	3905	846	PCPPAIMYQSSNKC  MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEQGSPSEMVHSSEFRVDNNHLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFIYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS. LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

are In	CEO TE	16-4	T CEC	T	<u> </u>	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
	1		USSN	nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-			location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ľ	ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
			ļ	sequence		nucleotide insertion
			}	İ		DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI
						QCQNKGWDGYDVQWECKTDLDIAYKFGKT
		ł		ĺ		VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL
						GLQKLKESGKQHGFASFSDYYYKWSSADSC
			l	Ì		NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP
				]	1	YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ
	ļ		'			NTGHGATSGFGSAFTGQQGYENSGPGFWTGL
	1	ŀ	1		}	GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY
			ļ	ļ		PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA
						SGYGGTRRR
957	2307	A	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL
		ļ				VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVO
			ļ	}		DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV
	1	1	1			PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF
	[		ſ		ĺ	QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA
				ļ		HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG
			ļ			TLVGLGLLAGLVLLYHRRGKALEEPANDIKE
				[		DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL
				ļ		
,		1		f .		RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG
	ļ	ļ .	ļ			AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS
958	2308	A	8161	2340	1192	QAGSLV
730	2308	^	0101	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV
		1			·	EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP
						LNATLVITFEITFRSKNITILELPDEVVVPPGVT
	ŀ	ł		}		NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR
			ļ			FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM
			ĺ			NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL
			J			LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH
		l				AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL
	1	}		Ì		AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL
			1			AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL
	İ	<b>!</b>	İ			DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK
		}	}			FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE
			1			MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI
			l			ASPDVKLNLGGDFIKESTATTFLRQRGYGWL
ĺ	ĺ					LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV
İ						LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS
i			1			MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG
l		1	1			GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF
	ł			·		EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK
ĺ	1		ĺ			KPLLIYPIFLLYIYFLSLYTGV
960	2310	Α	8167	1	2921	
200	1 2210	^	310/	1	2741	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN
		i	l .			LYSQLNALQFTVDERSILWLNQFLLDLKQSL NOFMAVYKLNDNSKSDEHVDVRVDGLMLK
i		Ī				I NI IRAZAVYKI ALINISKSTEUVINVOVDATIATV
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS
			!			FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP
		i	!			FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFORHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP
						FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY
						FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI
						FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI
						FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGOKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL
						FVIPSEVKSECHODOPRAISIOSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA
						FVIPSEVKSECHODOPRAISIQSSEMIATNTRH CPNCRHSDLEALFODFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEODTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSOFWIDYEGM KSGKGRPISFVDSFPLSIWICOPTRYAESOKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN
						FVIPSEVKSECHODOPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFFFSPSSSEADI HLVHVHKHVSMQNHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD
						FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DISKETPPVRTLKSQSSLSGKPKERCPPNLAP LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN AGANLQNYGETSPDAISTNSEGAQENHDDLM SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
961	2311	A	8172	1442	682	EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN FPEFSFDFTREQLMEENESLKQELAKAKMAL AEAHLEKDALLHHIKKMTVE TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI
				-		ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG EVQVSDKERHTQLEQMFRDIATIVADKCVNP ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS VLRRMQKKYWKTKQVFIKATGKKEDEHLVA SDAFLDAKLEVFHSVQETCTELLKIIEKYQLR LNGMKS
963	2313	Ą	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV AWLSRAEWDQVTVYLFCDDHKLQRYALNRI TVWRSRSGNELPLAVASTADLIRCKLLDVTG GLGTDELRLLYGMALVRFVNLISERKTKFAK VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLQELSAEERQKYLE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQQPKKK  RSFSLSFSLLSPSEMMALGAAGATRVFVAMV AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ
967	2317	A .	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i 	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Seduence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT
-		<u> </u>				YMDNWRQNTTAEQEALQAKVDAENFYYVIL
		ĺ				YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATIHENIG
		[				AAGFKMSP
969	2319	A	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR
}	10.17	1	-	_		VPGRLLLLLPALCCLPGAARAAAAAAGAGN
ŀ			Ì			RAAVAVAVARADEAEAPFAGQNWLKSYGY
1	ļ	]	ļ			LLPYDSRASALHSAKALQSAVSTMQQFYGIP
1		1	Ī			VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGOKWROKHITYSIHNYTPKVGELD
						TRKAIROAFDVWOKVTPLTFEEVPYHEIKSDR
	]	]				KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF
						PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL
						VAVHELGHALGLEHSSDPSAIMAPFYQYMET
	į					HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT
	ł	1				KPNICDGNFNTVALFRGEMFVFKDRWFWRL
	İ	ļ				RNNRVQEGYPMQIEQFWKGLPARIDAAYER
	}		l			ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG
						ELGSCLPREGIDTALRWEPVGKTYFFKGERY
						WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY
			İ			PRNILRDWMGCNQKEVERRKERRLPQDDVDI
[			ļ			MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI
L						FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN
			İ			DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN
				}		YELLEKEVEKRKVALQEAKLKAKGLNPDGTP
1		ł			1	ALSTLGGFSPASKPSSPREVKAEEKSPISINVK
1		}				TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR
		ļ				SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY
1		:				SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA
						KKHRHERGHHRDRRERSRSFERSHKSKHHGG
j		)	j		ļ	SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE
					I	RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEOESFFEGKNMALFEEEMDSNPMVSSLL
						NKLANYTNLSQGVVEHEEDEESRRREAKAPR
1		1				MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG
	1					VLESFLIVAMCCTCTMLTAISMSAIATNGVVP
1	Ì			1	1	AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA
				1		GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA
1		1	1	1		AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV
				l		CLLGNRTLSRRSFDACVKAYGIHNNSATSAL
	İ			1		WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA
		ļ	1	}		ASGVFLENLWSTYAHAGAFVEKKGVPSVPV
						AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG
1.		l	1			IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM
			[	1	1	LAWPSPWVIVIGSFFSTCGAGLQTLTGAPRLL
		-	· ·			QAIARDGIVPFLQVFGHGKANGEPTWALLLT
1	1	<u>†</u>	1	<u> </u>	•	VLICETGILIASLDSVAPILSMFFLMCYLFVNL
1			1	1		ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL
1	1	1	1			CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT
1		1	]			KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ
L	J	L	ــــــــــــــــــــــــــــــــــــــ	<del></del>	<u> </u>	ייייי ליייייי לייייייייייייייייייייייי

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ĺ	ĺ	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	}	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
	l			sequence	1	nucleotide insertion
		<del></del>	<del></del>	Sequence		LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE
				]		NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ
	Ì	}	ļ			SAGLGGLKHNTVLMAWPASWKOEDNPFSW
				ļ		KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ
				İ	į	ERFGGGHIDVWWIVHDGGMLMLLPFLLROH
	]	}	ł		ļ	KVWRKCRMRIFTVAQVDDNSIQMKKDLOMF
				ļ		LYHLRISAEVEVVEMVENDISAFTYERTLMM
		}		1	}	EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS
						HTAAAARTQAPPTPDKVQMTWTREKLIAEK
					]	YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV
	ļ	ļ			)	KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD
						ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV
	ł	<b> </b>			İ	RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ
						VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ
				ļ	1	REKANGTTVHVGIHPSKVVTTRLKLDKDRKKI
0.50	2222	<u> </u>				LERKAKSRQVGKEKGKYKEELIEKMQE
973	2323	A	8237	873	4610	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC
	İ	ĺ		•		PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG
		<b> </b>				AKAWIMDEEEDAEEEGAGGRQDPSRRSIRLR
				İ		PLPSPSPSAAAGGTESRSSALGAADSEGPARG
	ĺ	İ			ľ	AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG
					i	SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS
		1				CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG
	İ	İ		ĺ		QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA
			1	!		VEREQERVKSAGFWIIHPYSDFRFYWDLTML
					<u> </u>	LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD
	Ì	Ì	ĺ	ĺ		TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK
		]	}	j		YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK
	1			ļ		TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE
						EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG
	}	]	J	].		CLQFLVPMLQDFPDDCWVSINNMVNNSWGK
				•		QYSYALFKAMSHMLCIGYGRQAPVGMSDV
		ļ				WLTMLSMIVGATCYAMFIGHATALIQSLDSS
		}				RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD
	İ	•				YYEHRYQGKMFDEESILGELSEPLREEIINFNC
						RKLVASMPLFANADPNFVTSMLTKLRFEVFQ
	}	1		}		PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE
	1	1				TKLADGSYFGEICLLTRGRRTASVRADTYCR
	ł	1		1		LYSLSVDNFNEVLEEYPMMRRAFETVALDRL
		[				DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ
	1			1		QIVQHDREMAHCAHRVQAAASATPTPTPVIW
	1	}		}		TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR
						PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS
						ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS
	ļ	1				PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA
	1					GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ
						AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS
	1		}			SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP
	1	1				PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR
		1				RGTPPLTPGRLTQDLKLISASQPALPQDGAQT
		1	1			LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS
	ł	[				GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP
	1	ł	1	}		LSLFGARATSSGGPPLTAGPQREPGARPEPVR
						SKLPSNL
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL
	1 2227	1	027/	217	100	LVPVKDASRICSLTYLLGSHWNNLVVRSPVL
						G

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA P
976	2326	A	8257	298	7086	GNMACWPQLRLLLWKNLTFRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTFMENSGEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNILPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTIILSTHHMDEADVL LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEFLKVAEESGVDA ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

	600 -	·	OF C	Dec distant	I Don't it is a	
	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
	peptide		un USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		-		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence	1	09/496 914	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
<b>)</b>				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		-				LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
						NEQYTFVSNDAPEDTGTLELLNALTKDPGFG
						TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
	1					DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
						CPPGAGGLPPPQRKQNTADILQDLTGRNISDY
	ĺ	[				LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS
		1				NTQALPPSQEVNDATKQMKKHLKLAKDSSA
						DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
						HAISSFLNVINNAILRANLOKGENPSHYGITAF
	- 1	ļ				NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
1		1				MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
	}					YWLSNFVWDMCNYVVPATLVIIIFICFOOKSY
						VSSTNLPVLALLLLLYGWSITPLMYPASFVFK
						IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN
		- 1				KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ
		1				AMADALERFGENRFVSPLSWDLVGRNLFAM
						AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
				'		DEDEDVRRERQRILDGGGQNDILEIKELTKIY
}	-					RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
1		1				
						SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
						HQNMGYCPQFDAITELLTGREHVEFFALLRG
l						VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY
	ľ	- 1				SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
		-				PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
		1				EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
						GYTTVVRIAGSNPDLKPVQDFFGLAFPGSVPK
		J				EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH
					•	IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
						SLHKNQTVVDVAVLTSFLQDEKVKESYV
977	2327	Α	8260	3	1567	IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG
						YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
		1				TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
1		1				SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
						YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS
						LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM
]		- 1				VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
		1				MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
		ļ				MYAYAGWFYLNFVTEEVENPEKTIPLAICISM
		1				AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
		1				FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV
1 1	1	ł			į	SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
]	l					LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA
		1				VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC
						LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
		İ				WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
978	2328	A	8261	2	2165	RGGSLRCVLGKLLGOLLCFOSERCVRFPEGLL
	-520		0201	-		
						RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP
	]				, ,	LADAASMSGVRAVRISIESACEKQVHEVGLD
] ]	Ì	į				GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE
						EEAAGTEGDAQEWPGAGSSADQDDEEGVVK
						FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI
						VRDKKFMTLDPVSQDALPPKQNPQTLQLISK
		ļ				KKSLAGAAQILLKGAERLTKSVTENQENKLQ
1 1	- {	l				RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
į l						AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
1 3						DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
		I				
						LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
						FAQLSREAVQIKSQVPHIVVKNQIISOPFPSLO
						FAQLSREAVQIKSQVPHIVVKNQIISOPFPSLO
						LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND
						VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289		1053	FVWNPRGGRKRRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLPVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGSPGACPAL GTKSCSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	Α	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFIYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

CEO ID	SEQ ID	Mar	CEA	Dendist 4	Dendista	14-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-
SEQ ID NO: of	NO: of	Met	SEQ ID NO:	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	noa	1	beginning nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		in USSN			F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
	dence	1	914	correspondi	acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
ì		l	ĺ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	i	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		(		peptide		/=possible nucleotide deletion, \=possible
<u></u>	<del> </del>	<b> </b>		sequence		nucleotide insertion
]						EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV
1						RHVLSCLGGGLALWRAGQWLWAQRLGHCH
	Į	ļ				TYWAVSEELLPNSGHGPDGEVPKDKEGGVF
1				ĺ		DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC
1		١.	ł		i	VGESWPQDQPWTKRLVMVKVVPTCLRALVE
		1 '				MARVGGASSLENTVDLHISNSHPLSLTSDQY
	<u> </u>					KAYLQDLVEGMDFQGPGES
984	2334	Α	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY
ĺ		1				TIREVVTEIRDKATRRRLAVLPYELRFKEPLPE
		İ				YVRLVTEFSKKTGDYPSLSATDIQVLALTYOL
	ĺ					EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS
						GFHLPYKPKPPQETEKGHSACEPENLEFSSFM
1						FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE
	[	l			}	NGFEDRKDDSDDDGGGWITPSNIKQIQQELE
		1		İ		QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV
1		İ				LAVNGMLIREARSYILRCHGCFKTTSDMSRV
						FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP
						KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF
						PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI
1	ļ					SSRSATLQVRDSTLGAGRRRLNPNASRKKFV
1	İ	1	İ			KKR
985	2335	A	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET
1			55	552	325	EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE
100	2330	'`	0323	87	1172	I
ĺ	i	i				GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL
1			i			FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG
ļ	1					YYAADQWVFGLGLCKMISWMYLVGFYSGIF
Ì	ľ					FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS
	į.	į				LATWSVAVFASLPGFLFSTCYTERNHTYCKT
ł	i		i i			KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY
						SMIRTLQHCKNEKKNKAVKMIFAVVVLFLG
	1					FWTPYNIVLFLETLVELEVLQDCTFERYLDYA
	1	İ				IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL
		[				FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS
1005	0000	<b> </b>	0000			TMDHDLHDAL
987	2337	Α	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC
						GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN
		ĺ				VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD
		1				GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK
		<u></u>				LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	Α	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA
1	1	1				VAEVRLPSATLCYFCRCRLGLGAALFPRSAR
						ALAASALPAQGSRWPVLSSPGLPAAFASFPAC
						PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV
			<b> </b> .			RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH
-						VSKLLSQCKFDLLEELVAKEVLHALKEKVTS
İ	1	ł				LPDNHKNALAANIDEIVFTSTGDISIYYDEKG
						RKFVNILMCFWYLTSANIPSETLRGASVFQVK
1		l				LGNQNVETKQLLSASYEFQREFTQGVKPDWT
			]			IARIEHSKLLE .
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL
1	2000	1 1	0.77	۱ ،	107	KSLHPMS
990	2340	A	8361	210	1115	
""	2540	^ <b>.</b>	ا يوده ا	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ
İ		l				ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL
						DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF
1		ļ				FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW
						LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG
1					j	QKALLHYNEEAVQINPKCFYTPKCHQDRNDL
L	L	<u></u>				LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	delice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uchico			1 717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	boquento	/=possible nucleotide deletion, \=possible
				sequence	· ·	nucleotide insertion
ļ		<del></del>	<del> </del>	soquonoc		LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV
						RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH
J		ļ				OFKLCIMRRSKGRAEKS
991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL
751	2341	^	6505	^	121	EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT
		ļ			1	EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK
	1		ļ		ļ	GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK
		1				SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP
		ľ				HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE
						DATLENGCLWFIPGSHTSGVSRRMVRAPVGS
Į	}				İ	APGTSFLGSEPARDNSLFVPTPVQRGALVLIH
1.		ŀ				GEVVHKSKQNLSDRSRQAYTFHLMEASGTT
İ	Ì					WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906	4	MALSGNCSRYYPREQGSAVPNSFPEVVELNV
772	2542	^	6570	200	"	GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN
		]		j		DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV
					}	LPDHFPEKGRLKREAEYFOLPDLVKLLTPDEI
1	<u> </u>	ľ			ţ.	KOSPDEFCHSDFEDASQGSDTRICPPSSLLPAD
		Į.				RKWGFITVGYRGSCTLGREGQADAKFRRVPR
	}	1		ļ	•	ILVCGRISLAKEVFGETLNESRDPDRAPERYTS
1	1					RFYLKFKHLMGAPASNFILGFWGLGQNQDK
ſ		ŀ	j	ļ		HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA
	'					QEVSRRRRWLGDPEHL
993	2343	A	8379	1	2794	MRMQRHKNDTMDFGDSGKFIGGGVLCLLHQ
1 773	2545	^	6577	•	2//	SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE
ŀ				}		DMVTTASTYLFEATEKRFFFKNVSILIPENWK
ł	i	l		ł	į	ENPOYKRPKHENHKHADVIVAPPTLPGRDEP
						YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG
Į						PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA
1			1			KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA
1	}	1	·	ł		CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM
		l				QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST
İ				ł	İ	WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI
1			1	1	[	VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ
1						TVENGSWVGMVHFDSTATIVNKLIQIKSSDER
1						NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH
Ì	1		]	<u> </u>	<b>J</b>	SOLDGSEVLLLTDGEDNTASSCIDEVKQSGAI
1		1	ŀ			VHFIALGRAADEAVIEMSKITGGSHFYVSDEA
1		1	1		1	ONNGLIDAFGALTSGNTDLSQKSLQLESKGLT
	1	}	}	1	1	LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP
		]		1	1	SISLWDPSGTIMENFTVDATSKMAYLSIPGTA
	ļ					KVGTWAYNLQAKANPETLTITVTSRAANSSV
				1		PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP
1	1	1	}	l	1	VLGANVTAFIESONGHTEVLELLDNGAGADS
İ		ļ		]	1	FKNDGVYSRYFTAYTENGRYSLKVRAHGGA
				1	1	NTARLKLRPPLNRAAYIPGWVVNGEIEANPP
1	İ	[	1	1	1	RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL
			1	1	1	PLPDQYPPSQITDLDATVHEDKIILTWTAPGD
	1	1	1	1		NFDVGKVQRYIIRISASILDLRDSFDDALQVN
1	1		1		1	TTDLSPKEANSKESFAFKPENISEENATHIFIAI
	1	i	1	1	1	KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT
						PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
J	J	j .	1	]	1	NFILSTTI
994	2344	Ā	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP
	1	1	1		1	RONPGIFYWIFLPSRSHSASHGSRQRQVSCQG
			Ī			TQDEILKMRNTFAELKNSLEALSSRMDQAEE
1	]	1	l	}	Į.	RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
			1	1		PSSWDYRACLS
995	2345	A	8390	194	3421	AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR
				<del></del>		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		DFLSMKQSPALAPEERCRRAGSPKPVLRADD NNMGNGCSQKLATANLLRFLLLVLIPCICALV LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE
						QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPPPEQWLEPDTYCYITGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIOTFLLN
996	2346		8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSRSRSPSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
						ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ľ				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l		i	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
		ļ	<del> </del>	sequence		GTKDFVGPSERGGGRARGTFQFRARGRGWG
1		ļ	1	ļ		RGNYSGNNNNNSNNDFQKRNREEEWDPEYT
1						PKSKKYYLHDDREGEGSDKWVSRGRGRGAF
1		Ì				PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE
1		İ				DDESGTENREEKDNIQPTTE
997	2347	A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS
""	2341	Ι Δ	6576	202	332	KQENLESLEATGREEEGGQGPPVTTKGVLLA
1						LLMAGLALQPGTALLCYSCKAQVSNEDCLQ
Ì						VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN
1 330	2540	Α	0400	057	301	MSDPRRPNKVLRYKPPPSECNPALDDPTPDY
	,	l				MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI
1.		]		ļ	ļ	SFANSRSSEDTKOMMSSFMLSISAVVMSYLO
]		İ		1		NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL
***	,		""	~		LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL
}		1	1	1	ł	VSKKANQQLNFTEAKEACRLLGLSLAGKDQ
1						VETALKASFETCSYGWVGDGFVVISRISPNPK
j		]		j .		CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW
		İ				TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS
			1			VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE
			1			VFMETSTMSTETEPFVENKAAFKNEAAGFGG
1		1	}	}	Ī	VPTALLVLALLFFGAAAGLGFCYVKRYVKAF
		'				PFTNKNQQKEMIETKVVKEEKANDSNPNEES
\$		•				KKTDKNPEESKSPSKTTMRCLEAEV
1000	2350	A	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT
1000		1	0.00	-	i '''	AAIFTADGNMISASTLMDILLMNDFKLVINKI
				,		AYDVQCPKREKPSNEHTAEMEHMKSLVHRL
						FTILHLEESQKKREHHLLEKIDHLKEQLQPLE
}			j	1	ŀ	QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA
				ł	1	LAWLTWWVYSWDIMEPVTYFITFANSMVFF
		ł			ļ	AYFIVTRODYTYSAVKSRQFLQFFHKKSKQQ
					İ	HFDVQQYNKLKEDLAKAKESLKQARHSLCL
1	}	l			ł	OMOVEELNEKN
1001	2351	A	8410	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARG
		i				TGALLLRGSLLASGRAPRRASSGLPRNTVVLF
					<b>{</b>	VPQQEAWVVERMGRFHRILEPGLNILIPVLDR
	1	1		1	ļ	IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV
		l	ļ	ì		LYLRIMDPYKASYGVEDPEYAVTQLAQTTM
				1	1	RSELGKLSLDKVFRERESLNASIVDAINQAAD
				}		CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE
1	[	[		1		RRKRATVLESEGTRESAINVAEGKKQAQILAS
	ļ	1		1		EAEKAEQINQAAGEASAVLAKAKAKAEAIRI
				1		LAAALTQHNGDAAASLTVAEQYVSAFSKLA
J		ļ	]	1		KDSNTILLPSNPGDVTSMVAQAMGVYGALT
	İ	I	1			KAPVPGTPDSLSSGSSRDVOGTDASLDEELDR
				1		VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVOLRTTNECHKTY
		1		1 :	1	YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG
		İ	Ì		1	NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL
· ·				1		AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC
ŀ				[		PKCTQIINGSVDVDTEDRQKRKPESDGRTAK
1	1	l	1	1	ł	ALRSLQFTNPGRQTEFAPETGKREKRRLTKN
	]			1		ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC
İ	1			1		DCLDEDCLGCFYACPACGSTKCGAECRCDRK
1	1	l	ľ	ľ	1	WLYEQIEIEGGEIIHNKHAG
1003	2353	A	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR
1 - 3 - 3		1		1		CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT
1	ł	1	ļ	1		MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT
	l	1	1		<del></del>	

000 VV	100000	<u> </u>	Tara -		Table 1	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	j	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	!	]	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	ļ <u>.</u>		ļ	sequence		nucleotide insertion
		l	1			SLDTENIDEILNNADVALVNFYADWCRFSQM
	İ			ĺ		LHPIFEEASDVIKEEFPNENQVVFARVDCDQH
						SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ
		1	ļ			RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
	]	ļ	j	]	J	KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
						LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY
						LGAMTNFDVTYNWIQDKCVPLVREITFENGE
•	1	ł				ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL
		Ì	į			ISEKGTINFLHADCDKFRHPLLHIQKTPADCP
						VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
	ľ	Ì	İ	i	i	HSGKLHREFHHGPDPTDTAPGEQAQDVASSP
1001	10055	<b> </b>	0.55	-		PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
	1	1	(	i	1	ACAAARSPADQDRFICIYPAYLNNKKTIAEGR
	1	1	1			RIPISKAVENPTATEIQDVCSAVGLNVFLEKN
		İ	ł	1		KMYSREWNRDVQYRGRVRVQLKQEDGSLC
	[	İ	1	ĺ		LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA
			L	<u> </u>		DQSLQQGEGSKKGKGKKKK
1005	2355	A	8453	90	530	QSHETKMQSGTHWRVLGLCLLSVGVWGQD
				i		GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE
	1	ł	} .	ł		ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF
			1			SELEQSGYYVCYPRGSKPEDANFYLYLRARG
		<u> </u>				NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW
	i			ĺ		KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS
						LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC
	<u> </u>	<u> </u>				QLCIFN
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL .
		j				SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT
				i		VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM
		1				GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA
				)		FWWHNKGLALIFCILQSLALTWYSLSFIPFAR
						DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP
		1		ļ		PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS
						DPRWGCVGPSMPTSTCLPGAVEASTTKASLP
	<u></u>	<u></u>	<u> </u>			KCPVDSSLPTPEACFL
1009	2359	Α	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP
	1	}		J		NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH
		l				LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
				1		FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL
		ļ	Į	<b>.</b>		TKLLVHSSLVGSILSALSALVGFIILSVKQATL
		1		]		NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD
		1		1		CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL
	1	ĺ	1			RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
		L		L		HDCGYEELLTS
1010	2360	Α	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA
		1	}			HRVALCHLAGCQEQAAWYHTLOILFFLVSAY
	].					FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT
		I		1		LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL
		1		1	•	SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
		]				GTLETQFTCPFCNHEKSCDVKMDRARNTGVI
		1	1			SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG
	1	1	1	<b>!</b>		PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC
	1					RAGFOCO
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM
	1		i	ĺ		RMKYGGQEFWADLNAMNVYETTEFDQLRR
						LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SVIRLIEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKRKKEYMNFKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI QYEEVSNTVSI
1013	2363	A	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	A	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIEL KG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2300	A	8311		403	KWIFSOFVRIFORFITALFAGHRRCHMAFAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTIFKNLQTVNVDEN
1017	2367	A	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAK WADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNITTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

NO. of mucl-colide colide sequence   NO. of mucl-colide colide sequence   NO. of mucl-colide   of peptide   NO. of peptide   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
							De A crartic Acid FeGlutamic Acid
Sequence			nou			6	E-Phenydologiae G-Clycine H-Histidine
1025   2375   A   8546   2194   1707   TVERFANKEN GRAYSHITE JERGKAN SPECIAL SPECIAL PRINCIPLES   1078   1		1			1		
1026							
### ### ### ### ### ### ### ### ### ##	seq-	uence		1			M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence    V=Tyrosine, X=Ulnknown, *=Siop codon, , *=Sossible mucleotide tellon, \=possible mucleotide tilon, \=possible mucleotide insertion   V=Tyrosine, X=Ulnknown, *=Siop codon, , *=	uence			914		acid residue	
peptide   sequence	l	ľ	l	i	amino acid	of peptide	
peptide   sequence	1			1	residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
		İ	1		nentide		
1025						· ·	
WIVEFFANWSDCQSEAPIYADLSLKYNCTG		<del> </del>	<del> </del>	<del> </del>	Sequence	<del></del>	
LNFGKYDVGRYTDVSTRYKYSTSPLITKQLPT	İ	l	ł		1	ĺ	
LILFQGGKEAMRRPQIDKKGAVSWTFEEDPVAS		ł					
1025   2375   A   8546   2194   1707   TVSPFIKTMASI.KCSTVVCVICLERKYREG.   1025   2375   A   8546   2194   1707   TVSPFIKTMASI.KCSTVVCVICLERKYREG.   1026   CRYPVCSVVCPRKHEGO.PRETPVEKKIRS   ALPITKTVEPVENKDDDDSIADFILMSDEEEDR   VSLQNLKNIL.GESATLISLILNPHLRQIMVN   DQGEDKAKLMRAYMQEPLFVEPADCCLGIV   EPSQNESS     1026   2376   A   8547   1078   594   VGMEDPAVNLKVILLGHWILLTTWGCIVFSGS   YAWANFILALGYWAVAQRDSIDAISMFLGG   LLATIFIDIVHISIFYPRYSULTDTGSYGMAIL   SLLLKPLSCCFVYHMYRERGGELLVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELLVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELLVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELLVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGELVHTGFIG.   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SLLLKPLSCCFVYHMYRERGGEN   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGMAIL   SSQDRSAYQTIDSABAFADPFAYPGGSSCRILL   SSQDRSAYQTIDSABAFADPFAYPGGSSCRILL   SSQDRSAYQTIDSABAFADPFAYPGGSSCRILL   SSQDRSAYQTIDSABAFADPFAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPAYPGGSCRILL   SSQDRSAYQTIDSABAFADPFAYPAYPAYPAYBAFATATABFATAFATABAFATAFATABAFATABFATAFATA		ł	ł				
1025   2375   A   8546   2194   1707   TVSFFKIMASI KCSTVVCVICLERFKYRCPA CRYPYCSVVCFRRHERQCMPETRPVEKKIR ALPTRTVREVPENDDOBLADH-INSDEEEDR VSLQNLKNLGESATLRSLLINPHLRQLMYNL DQGEDKAKLMRAYMQEPLFVFFADCCLGIV PRSQNLESS   VGMELPAVNLKVILLGHWILLTWGCIVFSGS YAWANFITLALGWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYBYRSLDTIGRGVGMAH SLLLKPLSCCVYVMYRRGGGELVHTGFLG SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SSQDRSAYQTIDSAEAPAPFAVFEGRSQDAR GY   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK   SKDRDAVDKLLKREELPGFLQ SGKDKDAVDKLLKREELPGFLQ SGKDKDAVDKLLKREELPGFLQ SGKDKDAVDKLLKREELPGFLQ SGKDKDAVDKLLKREELPGFLQ VAGTTSSTEDLANGDAAVTNINGSLSL   STDRDAVSKLEFTLANGDAVTSSTEDLANGDAVTSTEDLA		1		Ì			LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN
1025   2375   A		1	1	ļ			VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS
CRYPYCSVVCPRIHKEQCNPETRPVEKKIRS   APPTETVEKPVEKNDDDDSIADEINSDEEEDR   VSLQNLKNLGESATLRSILLNPHLRQLMVNLD   QGEDKAKLMRAYMQEPLFVEFADCCLGIV   EPSQNESS   VGMELRAVNLEVILLGHWLLTTWGCIVFSGS   VGMELRAVNLEVILLGHWLLTTWGCIVFSGS   YAWANFITLALGVWAVAQRDSIDAISMFLGG   LILATIFLDIVHISIFYPRVSLITDTGRFOYGMAIL   SLLLKPI.SCCYVHMYRERGGELLVHTOFLG   SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR   GY   GYGYTUDSAEAPADPFAVPEGRSQDAR   GY   PSRYSGSEGSTOTILLETAMGIIDV   SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR   GY   SRYSGSEGSTOTILKGELKVI MEKELPGFLQ   SGKDKDAVDKILKDI.DANGDAQVDFSEFIVF   VAAITSACHKYFEKAGLK		{	i	(	l	ĺ	TPTTVSDGENKKDK
CRYPYCSVVCPRIHKEQCNPETRPVEKKIRS   APPTETVEKPVEKNDDDDSIADEINSDEEEDR   VSLQNLKNLGESATLRSILLNPHLRQLMVNLD   QGEDKAKLMRAYMQEPLFVEFADCCLGIV   EPSQNESS   VGMELRAVNLEVILLGHWLLTTWGCIVFSGS   VGMELRAVNLEVILLGHWLLTTWGCIVFSGS   YAWANFITLALGVWAVAQRDSIDAISMFLGG   LILATIFLDIVHISIFYPRVSLITDTGRFOYGMAIL   SLLLKPI.SCCYVHMYRERGGELLVHTOFLG   SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR   GY   GYGYTUDSAEAPADPFAVPEGRSQDAR   GY   PSRYSGSEGSTOTILLETAMGIIDV   SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR   GY   SRYSGSEGSTOTILKGELKVI MEKELPGFLQ   SGKDKDAVDKILKDI.DANGDAQVDFSEFIVF   VAAITSACHKYFEKAGLK	1025	2375	Á	8546	2194	1707	
ALPITATVKPVENKODDDSIADFILNSBEEBR   VSLQNLKNLGESATLRSLLIPHRQLMVNL   DQGEDKAKLMRAYMQEPLFVEFADCCLGIV   EPSQNEES   VSLQNLKNLGESATLRSLLIPHRQLMVNL   DQGEDKAKLMRAYMQEPLFVEFADCCLGIV   EPSQNEES   VGMELPAVNLKVILLGHWLLTTWGCIVFSGS   YAWANFIILALGVWAVAQRDSIDAISMFLGG   LLATIFLDIVHRISTYPRVSLITDGRFGVGMAIL   SLLLKPLSCCFVYHMYRERGGELLVHTGFLG   SSQDRSAYQTIDSAEAPADPFAVPGGRGDAM   GY   STATE	1025	23/3	^ `	05-10	2254	''''	
VSLQNLKNLGESATLRSLLINPHLRQLMVNL	ļ	1	1	i		ì	AL DEKEMBER DE DE LA DEL MODEREDE
DQGEDKAKLMRAYMQEPLFVEFADCCLGIV	i		Ì	İ			
EPSONES	l		1	ľ	ł	1	
1026	1						DQGEDKAKLMRAYMQEPLFVEFADCCLGIV
YAWANFITLALGWAYAQROSIDASIMFLGG	l		ļ				
YAWANFITLALGWAYAQROSIDASIMFLGG	1026	2376	A	8547	1078	594	
LLATTFLDIVHISTPYPRVSLTDTGRFGVGMAIL   SLILKPLSCCFVYHMYRGGGELLYHTGFLG    SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY			١			1	
SLLLKPLSCCFVYHMYRERGELLVHTGFLG     SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR     GY	1		[		[		I I ATTEI DIVHIGIEVODVGI TOTGOEGVGMAII
SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY	1					1	
1027   2377   A   8557   1   340   DFLGPASPQEEGGSESSTMTELETAMGMIIDV   FSRYSGSEGSTQTLTKGELKYLMEKELPGFLQ   SGKDKDAVDKLLKDLDANADAQVDFSEFIVF   VAAITSACHKYPEKAGLK     1028   2378   A   8569   20   963   KMAATLGPLGSWQQWRRCLSARDGSRRLLL   LLLLGSQGPQQVQAGQTFEYLKREHSLSKP   YQGEARRPCFLRDWELQVHFKHGQGKKNL   HGDGLAIWYTKDRMQPGPYGRMDKFVGLQ   VYDTYPNEEKQGRVFPYISAMVNOSLSY   DHERDGRPTELGGCTAVRNLHYDTFLVTRY   VKRHLTIMMIDGKHEWRDLEVPGVRLPRG   YYFGTSSITGDLSDNHDVISLKLFELTVERTPE   EEKLHRDVFLPSVDNMKI.PEMTAPLPPLSGL   ALFLLYFFSLVFSVFAIVIGIILYNKWQEQSRK   RFY     1029   2379   A   8572   1   578   AAAASHRSRARSRRRVSSGPAPRAQSSAG   RVASGLDSAPLCTMARALCRLPRGLWLLLA   HHLFMTTACQEANYGALLRELCLTQFQVDM   EAVGETLWCDWGTRFSVFLADCTWHMAE   KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR   AVRDPPGSLTYFIVVPITVTLLVTALVVWQS   KRTEGIV     1030   2380   A   8574   1352   372   DSSTVKGGSESRHLCLIPLKGKARTREASSG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWYTYRSGRLSWGSKG   SRTCGRTSLCTSAKSSWGSVEIPFSTYPWEL   AIVPNVRISWRRGHFHGQSSYSTRPPSIHKDY   VNRLFLNWTEGGESGTRISNILREDGSVYF   CRVELDTRRSGRQQLQSIKGTKLITQAVTTT   TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT   AIVPNVRISWRRGHFHGQSSTSRPSHKDY   VNRLFLNWTEGGESGFLRISNILREDGSVYF   CRVELDTRRSGRQQLQSIKGTKLITQAVTTT   TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT   AIVPNVRISWRGGFAIVYYLIGLCLLLLWWRRKG   SRAPSSDF   SRAP	1		1	i			1
1027   2377   A   8557   1   340   DFLGPASPQEEGGSESSIMTELETAMGMIDV FSRYSGSEGSTQTILTKGELKVIMEKELPGFLQ SIKDKADVELLLKULMEKELPGFLQ SIKDKADVELLUKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK     1028   2378   A   8569   20   963   KMAATLGPLGSWQWRRCLSARDGSRRLLL LLLGSQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKHGQGKKNL HGDGLAIWYTKGHOPGPVFGMMBFVGLG VFYDTYPNEEKQGER VFYPISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYYFGTSSITGDLSDNHDVISLLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RRY     1029   2379   A   8572   1   578   AAAASHRSRARSRRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTRSYRELADCTWHMAE KLGCFWPNAEVDRTFFLAVHGRYFRSCPISGR AVRDPFGSLLYFFIVVITIVILLYTALVVWQS KRTEGIV     1030   2380   A   8574   1352   372   DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THILTITQALROPLHRAPLLPGGLCWSPRPLEK NKAMGRPLLIPLLLLJOPAFLJPGGSTGSGP SYLYGVTQFKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSSFSTRPSIHKDY VRIELTNWTEGGEGFLRISNLRKEDGSVFF CRYELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGILCLLLLWWRRKG SRAPSSDF     1031   2381   A   8580   905   340   RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFTVAGGFAIVYYLJQKFHSRALYYKLAV EQLQSIFEAQEALAVYHLAV GELCGPRIVGPRHAPTINHYLKLIDRENFYDI VDAKLKPYSGSKSGGLLYVHSSRGGPFORW HLDEVFILLKDQQQIPVFKLSGENGDEVKKE     1032   2382   A   8593   2558   961   RRRPRLLFGAEPCEPRVGPRRADMGCSAKAR	į		]	1			
FSRYSGSEGSTQTLTKGELKVLMEKELPOFLQ   SGKDKDAVDKLLKDLDANGDAVDFSEFIVF   VAAITSACHKYPEKAGLK	į.		1			1 _	
SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF	1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV
SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF			1	i			FSRYSGSEGSTOTLTKGELKVLMEKELPGFLO
VAAITSACHKYFEKAGLK	1		ļ			l	
1028   2378   A   8569   20   963   KMAATLGPLGSWQQWRCLSARDGSRRLLL   LLLLGSGGGPQQVGAGQTTEYLKREHSLSKP	1		1				
LLLLGSGQGPQQVĞAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKIFELTVERTPE EEKLHRDVPLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAASSHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCOWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTTQALAQPLHRAPLLPGOLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQFKHLSASMGGSVEIPFSTYYPBEL ALVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLREDQSVYF CRVELDTRRSGRQQLQIKGTKLTITQAVTTI TTWRPSSTTTAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGITYCFFKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1000	0000	<del> </del>	05.00	20	063	
YQGEAPRPCFI.RDWELQVIFIKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMYNNGSLSY DHERDGRPTELGGCTAIVRNI.HYDTFLVIRY VKRHLITMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLLYFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGL WLLLA HHLPMTTACQEANYGALLRELCL.TQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYFIVVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGF SYLYGVTQFKLASSMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGTLRISNLRREDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLJQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENOBEVKE  1032 2382 A 8593 2558 961 RRRFRLLPGAEPCEPRVGPRRADMGCSAKAR	1028	2378	A	8569	20	963	
HÖDGLAIWYTKDRMOPGPVFGNMDKFVGLG VFVDTYPNEEKQGRVFPYISAMYNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLURY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKJELTVERTPE EEKLJRRDVFLPSVDNMKLPEMTAPLPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTOFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHAPLLPGQLCWSFRPLEK NKAMGRPLLPLLLLLLPPAFLOPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL ALVPNVRISWRRGHHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITIQAVTIT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLLQKFHSRALTYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLTVHSSRGGFPQRW HLDEVFLELKDGQQPVFKLISGERGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLPGAEPCEPRVGPRRADMGCSAKAR		l .	ļ	1	1	l	
VFVDTYPNEEKQQERVFPYISAMVNINGSLSY DHERDGRPTELGGCTAIVRNILHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 I 578 AAAASIRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSLYPPIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHEHGQSFYSTRPPSHKDY VNRLFLNWTEQQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTTQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMVQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFPQRW HLDEVFLELKDGQQPVFKLISGERGDEVKEE  1032 2382 A 8593 2558 961 RRRPRLIPGAEPCEPRVGPRRADMGCSAKAR	•	1	}	i			YQGEAPRPCFLRDWELQVHFKIHGQGKKNL
DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETI.WCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESSGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMVQKYAGSRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFYAGGFAIVYYTLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVELKDGQQUPVFKLSGENGDEVKKE	i .	1	1		1	Į.	HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG
DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETI.WCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESSGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMVQKYAGSRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFYAGGFAIVYYTLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVELKDGQQUPVFKLSGENGDEVKKE		1	İ			1	VFVDTYPNEEKQQERVFPYISAMVNNGSLSY
VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFILVFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWILLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETL WCDWGRTTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTIQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLASSMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGILCLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGTYPCFPKGRTRHALCSVVLLLLTQQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLEISKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRAADMGCSAKAR	í	1		İ	l	ĺ	,
TYPEGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIIL YNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGGESGFLRISNLRKEDQSVYF CRVELDTRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHESSWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHISRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLELKDGQQIPVFKLSGGNGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	i	Ī		1			
EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIIL YNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLISNLRKEDQSVYF CRYELDTRRSGRQQLQSIKGTKLTITQAVTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFYDI VDAKKLKIPVSGKSEGLLYVHSSRGGPFQRW HLDEVFLEILKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1			1		ļ	1
ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY  1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRIPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLIPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTI TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHTYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	l				1		
1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSRPPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTLAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR		ļ					1
1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTOPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGTYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLIDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR			i	i	į	1	
RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLIPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1		}			ļ	
HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSFRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGKSEGLLYVHSSRGGFPQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1029	2379	Α	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG
HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSFRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGKSEGLLYVHSSRGGFPQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	İ				<b>!</b>		RVASGLDSAPLCTMARALCRLPRRGLWLLLA
EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMVQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFPQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1			,		
KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLJQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	i	1	1	1	i		
AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV  1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THILTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	1		1	İ	
KRTEGIV	1	1	1	1	1		
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE	1	İ		1		1	1
SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	L		<u> </u>	<u> </u>	<u>i</u> _	L	
SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG
THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR		1	İ			1	SRTCGRRTSLCTSAKSSWTYRSGRLSWOSIKG
NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	I		]			1	
SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	1	1			
AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1		İ	1		1	
VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	ĺ	ſ	1			
CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1		1	1		
TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	ŀ	1			
TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	!	1	1		CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT
AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1		1		1	Į	TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT
SRAPSSDF  1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	1	1	[		
1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	1	1	1	1	
AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	102:	0201	<del>  _   _     _  </del>	0500	1005	1340	
FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1031	2381	A	מאכא	905	J <sup>340</sup>	
EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	]	1	}	1		1	
EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	(	1	1	į.	FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV
VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE  1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1		1.	1			EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1		1			1	VDAKLKIPVSGSKSEGLI VVHSSRGGPFORW
1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1	1	1	1		1	
	1022	2202	12	0502	2550	061	DDDDDI I DC V CDCCDD ACCOUNT AD ACCOUNT AD
	1032	2382	/A	دوره	2338	1 201	
	L	<u> </u>	<u> </u>	<u></u>	L	L	WAAGALGYAGLLCAYLGAYMIYMYPSLIKQ

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Į.		ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ			1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
ł		İ				VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
ł		l				TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
į			,			MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
ŀ			·			GERAFMNRTVGEIMWGYKDPLVNLINKYFP
1		i	1			GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI
ŀ			1			SRIHLVDKWNGLSKVDFWHSDQCNMINGTS
			1			GQMWPPFMTPESSLEFYSPEACRSMKLMYKE
1						SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
	}		Ì			CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL
į			1			AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
	1					KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
		ĺ				LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
	1					GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	A	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
1055	2505	1 **	0373	375	<i>'</i> 0'	FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	A	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ
1 .03 .	1 200.	**	1 0327	010	104	PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL
						VNMGDRTSMVQDPGSQAPTSWISESQVFQTT
1						EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
ŀ		i		•		GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
ĺ		İ				F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
j		ļ				SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
				·		SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI
ļ	ł					HVYKKNGVGKVGDQILLAIKGQKKKALIVG
1		l				HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
						KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV
1						TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
ł	İ	ł	1			VAYRPISASVLSRPEASRTGEGSTVFNGAQNG
	<u> </u>			,		VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
1	1					VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
						AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	Α	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
}		ļ				GMVAHINNSRLKAKGVGQHDNAQNFGNQSF
į		!				EELRAACLRKGELFEDPLFPAEPSSLGFKDLG
	1	1				PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
]	<u> </u>	1				CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
	1	J	]			QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
		]				PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
	-				,	GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
1		1		}		PQNLLRLLRKAVERSSLMGCSIEVTSDSELES
1	1	!				MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
	1	1				RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL
	1	1	]		_	TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
					,	RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
	1	1				VVCTCLVALMQKNWRHARQQGAQLQTIGFV
1		İ				LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI
		ĺ	[			FINSREVSSQLRLPPGEYIIIPSTFEPHRDADFL
	1					LRVFTEKHSESWELDEVNYAEQLQEEKVSED
}	}	l				DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
		1	]	]		MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
}		1				KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
1		ł	i			LNSYEMRLVIEKAGIKLNNKVMOVLVARYA
Î	ľ	1	[		-	DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
		l	]			GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS
						~ <del></del>

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI EAL
1038	2388	A	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF GGIETLRVPSELVWLPEIVLENNIDGQFGVAY DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV LISGLVLLAYFLPAQAGGGKCTVSINVLLAQT VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP
1042	2392	Α	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT LDMITSTDHVLEQDFWICFTFYSVKERQI
1043	2393	A	8688	359	17	GLKTRAPATPTFQREVLGPAKQDMQRRCPRI GLMTSLLKPIKRRWRDYKRWKSGGFTGESC HHADTLGDRGGLQGDHSELLQWQKRILRTE GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL NLALADLLFALTLPIWAASKVNGWIFGTFLC KVVSLLKEVNFYSGILLLACISVDRYLAIVHA TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR RTVYSSNVSPACYEDMGNNTANWRMLLRIL PQSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLYNLVLLADTLM RTQVIQETCERRNHIDRALDATEILGILHSCLN PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA
1047	2397	A	8741	673	924	QLIQGGRLIKHEMTKTASA ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL
1048	2398	Α	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR

ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	SEQ ID NO: of nucl-cotide seq-uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VAVPNGQPPPSCMLLGGGAGPPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPHHWDKVIVDGS DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME NAGVNFVVSGREQAQHINTDGFKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWKQPTGSDELKIGEWSGPNQPN SSTGAWDNOKGHPLLENQGNAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNOMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEVQPTSDELKINGSWSGRPDENTSSWSGPPPPPGNVRPSNSSWSGPQPA TPKDEEPSGWEEPSPQSISRKMDDDGTSAWG DPNSYNYKNNLWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPNSWNKQHQQQPPQQPPPPQ PEASGSWGGPPPPPGNVRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDDDGTSAWG DPNSYNYKNYLWWSSASKPVSGWG EGGGNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPNSWNKQHQQQPPQQPPPPQ PEASGSWGGPPPPPPGNVRPSNSSWSGPQPA TPKDEEPSGWEEPSPQSISRKMDDDGTSAWG DPNSYNYKNYLWLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFPNSCLSP GLFNVGPQLSQQQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFPNSCLSP GLFNVGPQLSQQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNYGLPDLQTKGPIPGYGSGF SSGGMDYGGTAGGSPYNGGRAPPSSLPROPTNK IGSKSNASWPPEPGGVPWKGIQNIDPESDP YVTPGSVLGGTATSPTVDTDHQLRDNTTTG SNSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTMSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKOEAKAOTALHMCVLGNTTLLAEF
PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT							FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA
, , , , , , , , , , , , , , , , , , ,	1049	2399	A	8748	200	1387	PHRMGSPAPLLPGDLLGGGSDSI  VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT

Dotation   Dotation	SEQ ID NO: of	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence   Section   Sec	nucl- eotide			in HSSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence		•	ļ	1	-		
residue of poptide sequence poptide sequ	uence	!		914			
psptide   /-possible nucleotide deletion,possible nucleotide insertion   nucleotide insertion   nucleotide insertion   nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide insertion   Nucleotide   Nucleotid	ľ		ĺ				
Incledité insertion						sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
SNIKJIGSTAVLYCLCI.RHELSYRGGLALLIL   MAAGACYAAGGLQYORINT.PSPAAAASP   MPLHITPI.GLLLILILYCLISGLSSVYTELLMKR     CREPILALQNILP.YTFGVULNILGHAGGSOP   GLEGFSGWAALVVLSQALNGLIMSAVMKH   GRSTRILFVYSCSLVYNAVJSAVLRICLTAN     GRSTRILFVYSCSLVYNAVJSAVLRICLTAN     A 8758   3   1660   WVSSMGFEELLEQVGGFGFFGLRNVALLALP     FILATILLIGLAMRLYYGSR   WVSSMGFEELLEQVGGFFFGLRNVALLALP     RVLLPH.HFL.PIFLAAVPAHR.CALPGAPANS   HQDVWLEAHLPREPDGTISSCLRFAYPQAL     NTTIGEBERSKELEDERATYPCSGGWEYDH     SEFSSTIATESQWDLVCEQKGLINRAASTFFWTGGWL     LGLASSASVSYVMEAITRILTGSALAGFTILLYAYVSTLV     LGLASSASVSYVMEAITRILTGSALAGFTIL     MPLEEWLDVEHRITVAGVLSSTFWTGGWL     LALVGYLRDWRWLLLAVTLPCAPGLISK     WVSSASWLTQHOWKEAHRYLHGARLINGR     PVCEDSFSGEAVSKVAAGERVVRRSYSLDLD     GLGINVYQTQLLFGAVELPSKLLVYLSVRYA     GRRITQAGTILGTALAFGTRILYGSMKSWS     TVLAVMGKAFSEAAFTTAYLFTSELYPTUL     QTGMGITALVGRIGGSLAPLAALDDGVWLS     LPKLTYGGIALLAAGTALLIPETRQAQLFETI     QDVERKSAFTSLGEEMPMKQVEVEHEV     QTGMGITALVGRIGGSLAPLAALDDGVWLS     LPKLTYGGIALLAAGTALLIPETRQAQLFETI     QDVERKSAFTSLGEEMPMKQVEVEHEV     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELCQCPLCQLDCGSREGLIAND     TGNLNSDPLLELAQTFSVRKWALRR     TGNLNSDRLSRATAALPTSCRYSSPGGS     TGNLNSDPLLELGATTSVRKWALRR     TGNLNSDPLLELGATTSVRKWALRR     TGNLNSDPLLELGATTSVRKWALRR     TGNLNSDRLSRATAALPTSCRYSSPGGS     TGNLNSDRLSRATAALPTSCRYSSPGGS     TGNLNSDRLSRATAALPTSCRYSSPCT     TGNLNSDRLSRATAALTSTRATATITICT     TGNLNSDRLSRATAALTSTRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGNLNSDRLSRATATITICT     TGN			ļ				
MAGACYAAGGIQVPGNTLPSPPPAAAS	<b> </b>	<del> </del>	<u> </u>		sequence		
NPLHIPPI,GLLLILILYCLISGLSSVYTELLMKR   QRPLPALQNELT-YFTGVULNIGAGGSOP   GLEGRSGWAALVVLSQAINGILMSAYMKE     GSSTRIFLVVSCSLVVNAW 1.SAVILRLQLTAA     FFLATLLIGLAMRLYYGSR     WVSSMGFEELEQVGGFGPFQLRNVALLALP     FFLATLLIGLAMRLYYGSR     WVSSMGFEELEQVGGFGPFQLRNVALLALP     FRATLLIGLAMRLYYGSR     RVLLPH.HFL.PIFLAAVPAHRCALPGAPANS     HQDWWLEAHLPREPDGTISSCLRFAYPQALP     NTTIGERGNSRGLEDEPATYPCSQGWEYDH     SEFSSTIATESQWDLVCCQKGLINRAASTFFWTGGVM     SEFSSTIATESQWDLVCCQKGLINRAASTFFWTGGVM     LALVGYLRDWRWLLLAVTLPCAPGIS.BV     WVSSASWSVYMFAITRITGSALAGFTWAYSTLV     LGLASAASVSVYMFAITRITGSALAGFTWAYSTLV     LGLASAASVSVYMFAITRITGSALAGFTWAYSTLV     LGLASAASVSVYMFAITRITGSALAGFTWAYSTLV     LGLASAASVSVYMFAITRITGSALAGFTWAYSTLV     LGLASAASVSVYMFAITRITGSALAGFTWAYSTLV     LALVGYLRDWRWLLLAVTLPCAPGIS.BV     VPESARWLLTQAGFWKEAHRYLHICARLINGR     PVCEDSFSGEAVSVAAAGERVVRRPSVLJD     GLGLNVYOTQLLFGAVELPSKLLVYLSYMT     GRHTQAGTLLGTALAFGTRLL VSDMKSVS     TVLAVMGKAPSEAAFTTAYLFTSELYPTVLR     QTGMGLTALVGRIGGSLAFTTAYLFTSELYPTVLR     QTGMGLTALVGRIGGSLAFTTAYLFTSELYPTVLR     QTGMGLTALVGRIGGSGETALAALLDGVWAY     GRLTVAGGGALAAGTALLPETRQAQLPETI     QDVERKSAPTSLQEEMPMKQVQN     ERTPVAVSSASPGSGEGDEETTODEVSSHTS     EEGGGVVKVEKELENTEQPVGGREVVEHET     TGNLNSDPLILELCQFLCQLDCGSREGLIAHV     YQHTAAVVSAKSTMCPVGGRALSSPGSIGL     HLLIISEDQRSNCAVGABPTSHATFNSEKLP     EVINMESLFTVHNSGPSSAEGKDIASSPPVY     AGILLVCNNCAAVKSLLEAQTPSVRKWALRR     QNEPLEVRIQRLERERTAKKSRRDNETPEERR     VRRMRDREAKELGCMTSVRKWALRR     QNEPLEVRIQRLERERTAKKSRRDNETPEERR     VRRMRDREAKELGCMTSVRKWALRR     QNEPLEVRIQRLERERTAKKSRRDNETPEERR     VRRMRDREAKELGCMTSVRKWALRR     QNEPLEVRIQRLERERTAKKSRRDNETPEERR     VRRMRDREAKELGCMTSVSKLKYFFRAVDTA     ASVEYPYRTRRSCLIGMAFFEGNISSSH     HGHGGGBRGGGRVARRGGLGKYPFGRGR     PREVNYPTRRSCLIGMAFFEGNISSSH     ASVERDMANAA ASVEYPYTHONWEVYSKLKKTFADDIAGR     ASVEYPYTHONWEVYSKLKKTFADDIAGR     PREVNYPTHONWEVYSKLKTAADVAMAYGOS     ASTGROUND HAND AND AGGREN AND AND AGGREN AND AND AGGREN AND AND AGGREN AND AND AGGREN AND AND AGGREN AND AGGREN AND AND AGGREN AND AGGREN AND AGGREN AND AGGREN AND AG	'						
CRLPLALQNEFLYTFGVILINLGHLAGGSVMKPG							MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
GSSITILFVYSCSLVYNAVLSAYLLRLQLTAA   FFLATLIGIAMRLYYGSR							QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP
International Principles   International Princ							GLLEGFSGWAALVVLSQALNGLLMSAVMKH
1050   2400   A   8758   3   1660   WVSSMGFEELLEQVGGGFGFOLRNVALLALE RVLLPLIFILIPLE, AVPAHRCALFGAPANFS HQDVWLEAHLPREPDGTLSSCLRFAYPQALP NTTLGERGSRGELEDEFATYPCGGWEYDH SEFSSTIATESQWDLVCEGGRAKAASTFFFA GVLVGAVAFGVLSDRFGRRELLLVAYVSTLV LGLASAASVSVVMFAITENLSGALGFTIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLEDWWWLLAAVFYLEAGULSUW VPESARWLLTQGHVKEAHRYLHCARLNGR PVCEDSFSQEAVSKVAAGERVVRPSYLDLF RTPRELRHISLCCVVVWFGVNTSYGISLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRITQAGTLLGTALAFGTHIVSSDMKSWS TVLAWMGKAFSEAAFTTAVLFTSELYPTUR QTGMGTALVGRLGGSLAVAKVAAGERVVRPSSUDLF RTPRELRHISLCVVVWFGVLALDFOWLSS LYKLTYGGIALLAAGTALLIPETRQAQLPETI QDVERKSAPTSLQEEBMFMKQVQN TONGMGTALVGRLGGSLAVSKVAAGERVVRPSSUDLF EEDGGVVKVEKELENTEQPVGGREVVEHEV TONLNSPDLELCQCPLCQLGSREQLAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHHSEDGRSNCAVCGARFATTSTMSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAVRKLLEAQTFSVRKWALRR QNEPLEVRLQRERTAKKSRRDNETPEEBE VRRMBDREAKRLQRMQETDGGRARRLQRDR EAMRLKRANFEPKKAAGRIFTSKWALRR QNEPLEVRLQRERTTAKKSRRDNETPEEBE VRRMBDREAKRLQRMGTDGGRARRLQRDR EAMRLKRANFEPKKAAGRIFTSKWALRR LEKMDMMRAOFGQDPSAMAALAAGMNFF QLPYSGVELDSOLLLGKMAFEGNSSCH LEKMDMMRAOFGGGPVSAMAALAAGMNFF QLPYSGVELDSOLLLGKMAFEGNSSCH LEKMDMMRAOFGGGPSAMAALAAGMNFF QLPYSGVELDSOLLLGKMAFEGNSSCH AGROWSTYNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVVSGYSSQFGGY PATGADVAFSVNBILLGDPMANNAMAYGSSI ASHGKDWYHRELHFFVSGARTAGMAFHSGYGA HGSSGFVVGGGRVAFFGGLGRYFFGRAAGRAGHT PATGATAG PLITY WYTHELHFACH GENEDIFKE PATGATAG	1					i	
RVLIPLHFILIPIFLAAVPAHRCALPGAPANFS HQDVWILEAHILPREPIGTILSCLRFAYPQALP NTTILGERGSRGELEDER FYCSGGWEYDH SEFSSTIATESQWDLVCEQKGUNRAASTIFFA GVLVGAVAFGVISBRFGRRILLVAYVSTIV LGLASAASVSYVMFATIRTI.TGSALAGFTIIV MPLELWILDVEHRTVAGVISTWGGWAL LALVGYLRDWRWILLAVTI.PCAPGILSLWW VPESARWILTQGHVKEAHRYLLIAVTI.PCAPGILSLWW VPESARWILTQGHVKEAHRYLLIAVTI.PCAPGILSLWW VPESARWILTQGHVKEAHRYLLIAVTI.PCAPGILSLWW VPESARWILTQGHVKEAHRYLLIAVTI.PCAPGILSLWW GLGILNVYTQTQLIFGAVUE.PSKILLYVI.SVRYAW GRRI.TQAGTI.LGTALAFGTRILLVSSDMKSWS TVLAVMGKAFSEAAFTTAVT.PSELYPTVLR QTGMGI.TALVGRI.GGSLAPLAALLDGWUS. LPKILTYGGILALAGTALLJETRQAQLETI QDVERKSAPTSI.QEEEMPMKQVQN  1051 2401 A 8759 515 1625 EERGGVVKVEKELENTEQFVGNEVVEHEV TGNLNSDPLLELCQCPL.QLDCGSRGDLAHV YQHTAAVVSAKSYMCPVCGARTSGALAHV YQHTAAVVSAKSYMCPVCGARTSGALAFT VROMANASSYMCPVCGARTSGALARV YQHTAAVVSAKSYMCPVCGARTSGALARV VQHTAAVVSAKSYMCPVCGRARSGGGGVARRE QNEPLEVRLQRILERETAKKSRRONETPEERR QNEPLEVRLQRILERETAKKKSRRONETPEERR VRRMBDRSAKRLQRMQETDEGRARRLGRR LEKMDMMRAGGGGDVSARGGGKVARRECONSSUH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGKVARREGGLGKPYGGGA ASLVFVFTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDTGYGGYSOFGGY PATGADVAFSVNELLLGDPMANVAMAYGSSI ASHGKDMYHKEL HEPVSVGGYSOFGGY PATGADVAFSVNELLLGDPMANVAMAYGSSI ASHGKDMYHKEL HEPVSVGYSVSGGYSOFGGY PATGADVAFSVNELLLGDPMANVAMAYGSSI ASHGKDMYHKEL HEPVSVGYNSVGYSOFGGY PATGADVAFSVNELLLGDPMANVAMAYGSSI ASHGKDMYHKEL HEPVSVGYNSSAYTIYLLAGAAFQPLIIY WITHLVR ATVRSDLSTHILLAYSGYKYYGMILSVITGL LFGSGGYVALAWTSSALMYVSILTGLLIGGIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIKE VLMCCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDWGQGGPHRQKDVKIVTVEKK VNELNRLEKTKVERFPDLAAEKECRDREER NEKKAOJOEMKKREKEERKERKEREDELBER	1050	2400	Δ	8758	2	1660	
HQDVWLEAHLPREPDGTLSSCLERAYPQALP NTTLGERGNSGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFA GVLVGAVAFGYLSDRFGRRELLUAYVSTLV LGLASSASVSVMATIRTLTGSALAGFTIIV MPLELEWLDVERRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTCAGGILSLWW VPESARWLLTQGHVKEAHRYLLHCARINGR PVCEDSFSQEAVSKVAGERVYRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNNYQTQLLFGAVELPSKLLVYLVSVRYA GRRLTQAGTLGTALAFGTRELLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSTLAVGRLGGSTLAVGRLGGSLAVGRVGAGTLGTALAFGTRELLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSTLAVGRLGGSTLAVGRLGGSLAVGTLAVGRLGGGTLAVGRLGGGTLAVGRLGGGTLAVGRLGGGTLAVGRLGGGTLAVGRLGGGTLAVGTLAVGRLGGGTLAVGRLGGGTLAVGRLGGGTLAVGTLAVGRLGGGTLAVGTLAVGRLGGGTLAVGRLGGGTLAVGTLAVGTLAVGTLAVGTLAVGTLAVGTLAVGT	1030	2400	^	6736	3	1000	RVI.I.PI.HFI.I.PIFI.A.A.VPAHD.CAI.DCADANICS
NTTLGEERQSKGELEDEPATVPCSQWEYDH SEFSSTIATSQWDJ.VCEQKGINRAASTFFRA GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAJSVSYVMATIRTLTGSALAGFTIIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWILLAVTLPCAFGILSI.WW VPESARWLITQGHVKEHYLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNNYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAPTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTATLLYSSDMKSWS TVLAVMGKAFSEAAPTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTATLLYSSDMKSWS TVLAVMGKAFSEAAPTLAYLFTSELVPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTATLYLFTSELVPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALAAGTATLYLFTSELVPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALAAGTATLYLFTSELVPTVLR TONLNSDPLLELCCCPLCQLDCGSREQLIAHY VQHTAAVVSKSYMCYGAALSSFGSLGR HLLHISEDQRSNCAVCGARLSSFGSLGR HLLHISEDQRSNCAVCGARLSSFGSLGR HLLHISEDQRSNCAVCGARLSTRATFNSEKLP EVINNESLPTVHNBGPSSAEGKDJAFSPPVYF AGILLVCNNCAAYKKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERR VRRMRDREAKRLQRNGCARLSSRSLH RAGNLKRAANETPEERRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSLH HLEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSLH ASSIGNAVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNNEVQYSROPALPY RAGADVARSVMHLLGDPMANVAMAYGSSY ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNNEVQYSROPALPY RODLNAPDJYTPTMATTTVVLLAGMALGJQK RSSFEVLGLCASTALVWVVMEVLALLGLYL ATVRSDLSTFPHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVSLRTAAL GPDSMGGPVRQRLQLVTLTLGAAAFQELIY WITFHLUY SLAKKTANDVQQGFHAYGCKAAAVPKWSHRTAPGIMVF SNIKKTANDVQQGFHAYGKYCFCREAR VRTTSSSVNSSAYTYMGKKYENDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVANSIGCGKMNINVVVTTPV SNLKKTANDMOQQGFHRQKDVKIVTVEKK VMEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQJGGMKKREKEEMEREER							
SEFSSTIATESQWDL/VCEOKGINRASTFFFAA GVLVGAVAFGYLSDREFILLUAYVITL LGLASAASVSYVMFAITRILTGSALAGFTIIV MPLELEWLDVEHRTVAGVLSSTFYTGGVML LALVGYLIRDWRWLLLAVTLPCAFGILSLWW VESARWLLTQGHVKEAHRYLLHCARINGR PVCEDSFSGAVSKVAGENVRPSYYJGLSLDVS GLGLNNYQTQLLFGA VERVARFYLLHCARINGR PVCEDSFSGAVSKVAGENVRPSYYJGLSLDVS GLGLNNYQTQLLFGA VERVARFYLLHCARINGR RTPLRHISLCCVVVWFGVNFSYYJGLSDVS GLGLNNYQTQLLFGA VERVARFYLLGYBLLVSDMKSWS TVLAVMGKAFSBAAFTTAYLFTSELLYTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSARTSLQEEMFKQVQN  1051 2401 A 8759 515 1625 ERTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TONLNSDPLLELQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPSSLGR HLLHISEDGDKSNCAVCGARTSHATINSEKLP EVLNMESLPTVHNEGPSSAEGKOHAFSPVYVP AGILLVCNNCAAYRKLLEAQTPSVRK WALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMDREAKRLQRMQETDEGARRLGRDFP EAMRLKRANETPEKRQARLGRERAKLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEQNSSLH 1052 2402 A 8763 1106 70 RIGHGGRDRGGGRVARPGGLGRYFGRGAA ASLVFVPTRRSGPSCSARAAAAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNNEVQYSRDAPLPP RQDLNAPDLYPTMAGPTVLAGMALGJOK RSSEVLGLCASTALVWVVMEVLALLLGLVL ATVRSDLSTFHLLAYSGYKVVGRDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLITHLUS  1053 2403 A 8768 2 712 RPFRWYPELSELSAAAPRWSHRTAGINNYT VTYSSNSNSAATTYMGROK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLIMBCERTKVERFPDLAABKECRDREER NEKKANJGGCKMKNRVVNVYTTPV SNLKKTALDMOVQGJEHRGKDVKIVTTVEKK VNEILNRLEKTKVERFPDLAABKECRDREER NEKKANJGGCKMKTEKEEREMKKREMDELRSY							
LGIASAASVSYVMFAITRILTGSALAGFTIIV MPLELEVLIDVEHRIVAGVLSSTFWTGGVML LALVGYLIRDWRWILLAVTLPCAFGILSLWW VPESARWILTQGHVKEAHRYLHCARINGR PVCEDSFSQEAVSKVAGERVVRRPSYLDLF RTPRIRHISLCCVVWFGVNFSYYGLSLDVS GLGINVVQTQLLFGAVEPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGSAPLAALLDGWVLS LPKLTVGGIALLAGTALLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 ERTPVAVSSAPSGDSEDDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLIELCQCPLCQLDGGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHSEDQRSNCAVCGARPTSHATFNSEKLP EVINMESLFTVINEGPSSAEGKDIAFSPVYP AGILLVCNNCAAYRKLLEAQTFSVRKWALRR QNPPLEVRLQRLERERAKKSRRDNSTPERER VRRNRDREAKRLQRMGETDEGRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLCKMAFFEGNSSSLH RHGGBDRRGGGRARFGGLGRYPGGGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRRAAAPDPPPLFDTSGYSSQPGGY PATGADVAFSVNELLLGDPMANVAMAYGSSI ASHGKDMYHKELHRRVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQVSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVOMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGMAAFGPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFISSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAGOGEMKKREKERMELERSY							SEFSSTIATESQWDLVCEQKGLNRAASTFFFA
MPLELEWLDVEHRTVAGVILSTFWTGGVML LALVGYLRDWRWLLLAVTLPCAPGILSI WW VPESARWILTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVAAGGRVVRRPSYLDLF RTPRI.RHISLCCVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRITOAGTILLGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVQRLGGSLAPLAALLDGVWL QDVERKSAPTSLQEEEMPMKQVQN LFKLTYGGIALLAAGTTALLJETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN EERDGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLELCQCPLCQLDCGGREQLIAM YQHTAAVVSAKSYMCPVGGRALSSPGSLGR HLLHISEDQRSNCAVCGARFTSHATTNSEKLP EVINMESLPTVHNEGPSSAGGDLAFSPVYP AGILLVCNNCAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRIERERTAKKSRRDNETPEERE VRRMRDREAKRLGRMQETDEGRARRLGRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LERMDMMLRAQFGOSLAGMAHAAMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH TGSKHRARAAPDPPTLFDDTSGYSSQPGGY PATGADVAFSVNIELLGDPMANVAMYGSSI ASHGKDMYHKELHREVSVSKLKYFFAVDTA AVAKKLGLLVFPYTHQNWEVQVSRDAPLPP RQDLNAPDLYIFTMATTTYULLAGMALGIQK RFSPEVLGLCASTALVWVWEVQVSRDAPLPP RQDLNAPDLYIFTMATTTYULLAGMALGIQK RFSPEVLGLCASTALVWWWEVQLSLLGIJLLGLPL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAFGIMVF YFISSVNSSAYTIYMGKDKYRNENDLIKHGW PEDIWFHVDKLSSAHYYRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPEK SNLKKTADMDVGQIGFHRQKDVKKVTVTEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAGOGEMKKREEMKKKREMBELRSY	1						GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV
LALVGYLIRDWRWLLLAVTI.PCAPGILSI WW VPESARWILTQOHVKEAHRYLLHCARI.NGR PVCEDSPSQEAVSKVAAGERVVTRPSYILDIF RTPRI.RHISI.CCVVVWFGOVIPSYYGL SILDVS GLGI.NVQTQILJFGAVELPSKILLVYLSVRYA GRRITQAGTILJGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRIGGSLAPLAALLDGWWLS LPKLTYGGIALLAAGTALLJETRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGREVVEHEV TGNLNSDPLLELCQCPL.CQLDCGSRQLI.AHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILLIHSEDQRSIACOGARFTSHATFINSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILLAGTTSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMEDREAKRI.QTROGTEGQRARILLGRDR EAMRILKANNETPEKQARILIFEREAKRI.KRR LEKMDMMI.RAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNISSLH RIGHGGRDRRGGRVAPFGGLGRYPGGGAA ASLVFVPTRRSGPSGTASVAAMAYHSGVGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLIGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKILKYFFAVDTA YVAKKKIGLI.VPPHQNWEVQYSDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGICASTALVWVVMEVLALLLGLYL ATVRSDLSTHILLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQOCKMNNIVNVYYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNELINRLEKTKVERFFDLAAEKECRDREER NEKKAQOIGMKKREEMIKKREEMIELRSP					•		LGLASAASVSYVMFAITRTLTGSALAGFTIIV
VPESARWILITQGHYKEAHRYILHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYIDLF RTPRIRHISLCCVVVWGVINFSYYGLSIDVS GLGINYYQTQLLFGAVELPSKLLVYLSVRYA GRRITQAGTILGITALAFGTRILIVSSDMKSWS TVLAAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAFLAALLDGWILS LPKILTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELNTEQPVGGNEVVEHEV TGNLNSDPILELCQCPLCQDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHSEDQRSNCAVCGARFTSHATFNSEKLP EVINMESLPTVHNEGPSSAGKDIAFSPPVYP AGILLVCNNCAAYRKILLEAQTTSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLIGKMAFEEQNSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHPTRRSGPSGTASVAAMAYHSGYGA HGSKHPTRRSGPSGTASVAAMAYHSGYGA HGSKHPTRRSGPSGTASVAAMAYHSGYGA HGSKHPTRRSGPSGTASVAAMAYHSGYGA HGSKHPTRAGPPPLFDDTTSGGYSSQPGGYP PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTMAFITYVLLAGMALGJQK RFSFEVLGLCASTAUVVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYFQRLQLYTLTGAAAFQPLITY VLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELESAAAPRWSHRTAFGIMVF YFTSSSVNSSAYTITMGKDKYSREDLIKHGENIEDIPKE VLMMCSHYPTRAGGGCKMNNVVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAABEECRDREER NEKKAQIGEMKKREEMEKKREEMKKKREEMELKREY							
PVCEDSFSQRAVSKVAAGERVYRRPSYLDLF RTPRLRHISLCVVWGVNFSYYGJSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSDMKSWS TVLAVMGKAFSEAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVVERKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEEFTTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILHISEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTFSVRKWALRR QNEFLEVRLORLERETTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARFGGLGRYPGRGAA ASLVFVFTRRSGFSGTASVAAMAYHISGVGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGVGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGVGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGVGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMYHSGVGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRYSVSKLKVFFAVDTA YVAKKLGLLVFPYTHRQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYTTLGAAAFQPLIIY WLTTHLVR  1053 2403 A 8768 2 712 RPPRVWYPEIRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLIKGENIEDIPKE VLMMCAHLVKANSIQGCKMNIVVVTYPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNELINRLEKTKVERFPDLAAEECRDREER NEKKAQOIGMKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKKREEMEKEREMEELES							VPESARWILTOGHVKEAHRYIIHCARINGR
RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS							PVCEDSFSOEAVSKVAAGERVVRRPSYLDLF
GLGLNYYQTQLLFGAVELPSKILVYLSVRYA GRRITQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTTQDEVSSHTS EEDGGVVK VEKLEENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAAVSAKSYMCPVCGRALSSPGSLGR HILLHSEDQRSNCAVCGARTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYKKLLEAQTPSVRKWALRR QNEPLEVRLQRLEREERTAKKSRRDNETPEERE VRRMRDREAKRLQRMGETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNISSLH 1052 2402 A 8763 1106 70 RHGHGGRDRAGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSGI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTMAFITYVLLAGMALGIQK RPSPEVLGLCASTALVWVNMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSGGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLUR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNBLINRLEKTKVERFPDLAAEKCERDREER	,						RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS
TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN  1051 2401 A 8759 515 1625 ERTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TONLNSDPLLELCQCPLCQLDCGSREQLIAHY YQHTAAVVSASYMCPVCGRALSSPGSLGR HLLHSEDQRSNCAVCGARFTSHATTNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGGILLVONLOAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMPLKRAAAYRLJELAQTPSVRKWALRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFFEKONSSLH RSHGHGGRDVARFGGLGRYPGGGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQVSRDAPLPP RQDLNAPDLYPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVSLRSTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAAMYTLRLHKGENIEDIPKE VLMDCAHLJVKANSIQGCKMNVNVVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFDLAAAEKCERDREER NEKKAQIOEMKKREKEEMKKKERMDELRSY							GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA
OTGMGLTALVGRLGGSLAPLAALDGVWLS   LPKLTYGGIALLAAGTALLLPETRQAQLPETI							GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS
LPKLTYGGIALLAGTALLIPETROAQLPETI QDVERKSAPTSLQEEMPMKQVQN  1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVSAKSYMCPVGRALSSPGSLGH HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVPLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETTPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGRDRRGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHILLGDPMANVAMYGSGI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTTHLUX 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VUMDCAHLVKANSIQGCKMNNVNVYYTEW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR
DOST   2401   A   8759   515   1625   EIRTVAVSSAPSGDEGEETTQDEVSSHTS   EEDGGVVVVKELENTEQPVGGNEVVEHEV   TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV   YQHTAAVVSAKSYMCPVCGRALSSFOSLGR   HLLIHSEDQRSNCAVCGARFTSHATTNSEKLP   EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP   AGILLVCNNCAAYRKLLEAQITSVRKWALRR   QNEPLEVRLQRLERERFAKKSRR DEFTEERE   VRRMRDREAKRLQRMQETDEQRARRLQRDR   EAMRLKRANETPEERQARLIREREAKRLKRR   LEKMDMM RAOFGOPSAMAALAAEMNIF   QLPVSGVELDSQLLGKMAFEEQNSSSLH   RHGHGGRDRRGGGRVARPGGLGRYPGRGAA   ASLVFVPTRRSGPPSGTASVAAMAYHSGYGA   HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY   PATGADVAFSVNEILLGDPMANVAMAYGSSI   ASHGKDMVHKELHRFVSVSKLKYFFAVDTA   YVAKKLGLVFPYTHQNWEVQXSRDAPLPP   RQDLNAPDLYIPTMAFITYVLLAGMALGIQK   RFSPEVLGLASTALVWVVMEVLALLLGLYL   ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL   LFGSDGYYVALAWTSSALMYFIVRSLRTAAL   GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY   WLTHLUS							I PKI TYGGIALI AAGTALLI PETPOAGI PETI
1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGREVVEHEV TGNINSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHHSEDQRSNCAVCGARFTSHATFNSEKLP EVINNESHTYHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQAARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKOMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPMAFITYVLLAGMALGIQK RFSTENLLGLYLLGLYLLGLYLLGLYLLGLYLLGLYLLGLYLLG							
EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPILLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATTPNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNFPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGGAA ASLVFVPTTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLKTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIGEMKKREKEEMKKKREMDELRSY	1051	2401	Α	8759	515	1625	EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS
VQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATTNSEKLJ EVLNMESLPTVIHNEGPSSAEGKDIAFSPPVYP AGILL VCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGYELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTTHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVNSSAYTIYMGKDKYENEDLIKHGW PEDLWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							EEDGGVVKVEKELENTEQPVGGNEVVEHEV
HILIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLFITVHNEGFSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETFEKRQARLIREREAKRLKRR LEKMDMMIRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY	1						TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV
EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDM EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVWEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY							
AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVWMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							
QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							
VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMINF QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIOEMKKREKEEMKKKREMDELRSY							ONEPLEVRLORLERERTAKKSRRDNETPEERE
EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARPGGGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							VRRMRDREAKRLQRMQETDEQRARRLQRDR
QLPVSGVELDSQLLGKMAFEEQNSSSLH  1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1						EAMRLKRANETPEKRQARLIREREAKRLKRR
1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1		İ				
ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1052	2402	Α.	8763	1106	70	QLPVSGVELDSQLLGKMAFEEQNSSSLH
HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1032	2402	Λ.	8705	1100	, o	
PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							
ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							PATGADVAFSVNHLLGDPMANVAMAYGSSI
RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1 1				į	1	ASHGKDMVHKELHRFVSVSKLKYFFAVDTA
RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY							
ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1 1				ł	l	
LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY					ļ		
GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY		J		}		ļ	LFGSDGYYVALAWTSSALMVFIVRSLRTAAT
WLTFHLVR  1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY						ĺ	GPDSMGGPVPRQRLQLYLTLGAAAFOPLIIY
YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	I						WLTFHLVR
PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1053	2403	Α	8768	2	712	
VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	1 1	ł	l			1	
SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY	'		İ	j			PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE
VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY		j					VLIVID CARL V KANSIQGCKMNNVNVVYTPW
NEKKAQIQEMKKREKEEMKKKREMDELRSY							VNEILNRLEKTKVERFPDI AAFKECEDDEED
COLLECTION CONTOR CONTOR			ļ	-			NEKKAQIQEMKKREKEEMKKKREMDELRSY
55LMK VENMSSNQUGNDSDEFM							SSLMKVENMSSNQDGNDSDEFM
1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1054	2404	A	8769	344	527	REATTLACRNSCWVFSRCSLGACKPTVCSMP
SLSRQGSQTLCLRLAEYCMESVDSQRLLLS	L						SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ı	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł	ł	1	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide	1	/=possible nucleotide deletion, \=possible
ŀ			į.	sequence	1	nucleotide insertion
1055	2405	A	8770	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK
		[				KMLKCVVVGDGAVGKTCLLMSYANDAFPEE
·		1	1			YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ
ľ		l				EDYNOLRPLSYPNTDVFLICFSVVNPASYHNV
1			ì			QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK
}	ļ	1	1	ì	}	TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL
1	ŀ	l				ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS
		1				EGHSCCSII
1056	0406	<u> </u>	0.772	261	222	
1056	2406	A	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778 .	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH
ı	ŀ				ļ	RRDQKWHDKQYKKAHLGTALKANPFGGAS
1		ļ	1			HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK
1	1					NGKKITAFVPNDGCLNFIEENDEVLVAGFGR
ì	ļ	ļ		}		KGHAVGDIPGVRFKVVKVANVSLLALYKGK
1		l		<b>;</b>	l	KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL
1	l					VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME
		1			Į	TOSEPSELELDDVVITNPHIEAILENEDWIEDA
1	1		l	1		SGLMSHCIAILKICHTLTEKLVAMTMGSGAK
		1		}		MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
1						DPKLLDARTTALLLSVSHLVLVTRNACHLTG
	1			ţ		GLDWIDQSLSAAEEHLEVLREAALASEPDKG
1			1	1		LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLOGAIFVLLPHLGPILVWLFTRDHMSGWC
10.7	2407	'`	1 0005	2.10	1 '3'	EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY
1	1	1	l	1		LVWKDLGGGLGWPLALPLGLYAVQLTISWT
	ļ			İ		VLVLFFTVHNPGLALLHLLLLYGLVVSTALI
Į.	1	1	1	ł		WHPINKLAALLLLPYLAWLTVTSALTYHLWR
		1		ļ		DSLCPVHQPQPTEKSD
1000	0410	<del> </del>	0010	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1060	2410	A	8810			
1061	2411	Α	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA
	1	1	1		1	FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF
	i		i			GAMLNIAAVLCIATIYVRYKQVHALSPEENVI
		i	1			IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA
			1			HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH
1	1		ļ	ļ		GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS
1			İ	İ		GNFGTDLEQKLHWNPEDKGYVLHMITTAAE
	}		Į.	1		WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL
	<u> </u>	<u> </u>			<u></u>	TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824	1	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA
	1	i	1	1		GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA
ļ	1	1		!		YSLAPATPEVKVACSEDVDLPCTAPWDPQVP
1	]	1	1			YTVSWVKLLEGGEERMETPQEDHLRGQHYH
			}			QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR
1		1	1	1		CTLQDPDGQRNLSGKVILRVTGCPAQRKEET
	1				1	FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI
Į.	1					FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT
i	1		1			ELV
1063	2413	A.	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE
1005	1 2-13	1 **	1 3320	1 ***	1 32.	HORLHAWPPGPRRPSLWPPKNGKWHSGKRT
	1			1		AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ
					1	KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH
			1			CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL
1	]	1	]	1		TR
1000	1000	<del>  </del>	0025	2002	1960	
1064	2414	-A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE
						LNKQVSELSQLYKEAQAELEDYRKRKSLEDV
			1			TAEYIHKAEHEKLMQLTNVSRAKAEDALSE
}	1		1			MKSQYSKVLNELTQLKQLVDAQKENSVSITE
L			L		l	HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

ſ	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
-	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
-	nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
ĺ	cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
-[	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
1	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
-[					peptide	-	/=possible nucleotide deletion, \=possible
	1				sequence		nucleotide insertion
Ì							VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
١							SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
-[							EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
-							QAQEELAEMKRYSESSSKLEEDKDKKINEMS
1						1	KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
1							LQQQVKQLQNQLAECKKQHQEVISVYRMHL
J				]			LYAVQGQMDEDVQKVLKQILTMCKNQSQK
							K
ſ	1065	2415	Α	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
١	.						APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL
1							KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
1							TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
							DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1							DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
1	1000	0416		00.52	2006		AVLFITGIILTSGKCRQLSRLCRNHCR
	1066	2416	A	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG
ļ						!	RRRRGRVVSRKKMSLKSERRGIHVDQSDLL
							CKKGCGYYGNPAWQGFCSKCWREEYHKAR
							QKQIQEDWELAERLQREEEEAFASSQSSQGA
							QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
1						1	VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE
							FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE EQSECAQDFYHNVAERMQTRGKVPPERVEKI
							MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
-							QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV
-							KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
1							KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
-							QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
-				,			KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
- [							WSPDACLGVKQMYKNLDLLSQLNERQERIM
١							NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
1	ł						KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
ŀ	1067	2417	Α	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1							LRQAWATKQDPISKKK
Ī	1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
ţ	1069	2419	A	8857	1530	1583	PCRPGMECNSMISVHCNL
Ì	1070	2420	A	8866	293	1675	PYPQGGYPQGPYPQGGYPQGP
1							YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
j							EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
-							LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
							WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
-							VALSVLTASLSYMVGMIASFYNTEAVIMAVG
-						,	ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
-							VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
J							VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI
	[						FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
							WHGSASCTSPLSCPQAQPREKDASLQPSCMY
1							TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
١							HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
١							EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
1	105-					<u> </u>	GDMRSGGLIPVLSPE
1	1071	2421	Α	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
	-						DDKMGSNTFFKRNDCRYVMISCKADMAYDN
							VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII
ļ	1055	- 10:					LNGEKLKVFPVRSGT*QGCSVWP
1	1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
1			-				GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH
- 1							I (46-16-28-11-16-18-18-18-18-18-18-18-18-18-18-18-18-18-
١				1			LGRDWSWEKQKEQPKEYQRILQCFLDRKDC

CERCID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	liou	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1			USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		09/496		corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	i	914	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	1	ł	914	ng to first amino acid		T=Threonine, V=Valine, W=Tryptophan,
1				residue of	of peptide	
ł		1			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ				peptide	ł	/=possible nucleotide deletion, \=possible
	<u> </u>			sequence		nucleotide insertion
						CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ
						GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA
		L				RFPKKLCRVLPL\SADTAGLTGP
1073	2423	Α	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
			i			*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL
						RTSRYLPT*SIKSLGPDEPQEG
1074	2424	Α	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
i	į	į			•	KEISFGDYICHTFQGDCWADRSPLHEAAAHG
ł						RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
		1				*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	Α	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
						SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
1		}	j	}	j.	PWPSLLDKEREESLRQKRLSERERIGELGAPE
		İ				VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
						TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
į						EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
	1		-			QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
ł		ļ				QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
ļ		1	Į		ļ	PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR
	ļ					MEAVRTAKREPESTVLMRREPLHPFNPRRET
						KERE
1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
10.0	7.20	J	5022		1	*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
İ	Ì			į		FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
l		l				VSYTVSSRREVKAESSGKSHRGCPSPSPSEVR
i	ł	1	ł		ł	VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
						LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
1						RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
10,,	2427	1 ^	0,01	552	"	QYPALHRAGTEWQLSALHRAPRSTQPDKAC
					1	RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
ļ		l		}		\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
1078	2420	A	6505	230	/01	PALPFAATPGSRGQALCRGGRRRQHLHGPLH
				1		RP*OAAPALHAGCOLAPHPPT
1070	2420	-	9012	121	276	
1079	2429	A	8912	121	376 ·	NLIWKLCVTERRLVILDNYDLASE/YEANKYI
1		1				CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
1000		<u> </u>	0000	L	1500	EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
		•		1	1	GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
		Į				DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
1				l		YAPICMEYGRVTLPCRRLCQRAYSECSKLME
[					[	MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
1					İ	GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
1	1	1		}		SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
1				1		IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
1	ļ	1				WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA
		i	1			STVTQGSHNKACTMLFMILYFFTMAGSVWW
]		J	}	ļ		VILTITWFLAAVPKWGSEAIEKKALLFHASA
		l	1			WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
1						VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
1		ł		}	}	VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
1			1		i	VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
		1		~~		CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
		1				TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1		ŀ				LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	A	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV
1702	L432	1 4.7	10720	L 333	71017	1 1 011 001121 TATAKODIMKOOKKO YKKKY Y

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
}	}	•		sequence		nucleotide insertion
						GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
1						QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
1						FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
İ						EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK   HNNQILKTSYA*HQQS/RQIQKKMMEIMT*EV
						OTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH
1	[					DVFIRKVKMLENPGFER\MELRGGGSSS
1083	2433	A	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
		, 	•			WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
}						AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
1084	2434	A	8950	156	318	GWLLT/ESCSISPKLCSIAVH*DNPAWF HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
100.	2131	**	6,50	150	3.0	WGGKRVQPFWKRVWQKRTLNLRV
1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
						*TIYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR
						MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
				_		HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW   YSHKYY
1086	2436	A	8962	868	1026	H*KILQVGRAQRAHXSRL*SOLLRRLRHESHL
						NPGARGCSEARLHRCTPAWTT
1087	2437	A	8985	58	330	LHVKHLGHFQLVFSEVICHCILMPVS*ELQRL
		1				*ERSVCAFHVCIQTYVCLQVYACMCVYYICM
1088	2438	Ā	8989	394	404	FVYSVYGCGLCTCVCMDVYICVCVQEFL
1000	2438	Α.	0707	394	404	N*KWILHVNVRIQSIFF/IKRNQK/INSHELKLD KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA
						KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV
						YKDLSKLNRRKTE/S*/VKKWVKDLSRYFIKE
1000	2420		0001		200	VISMENKHKKIFSTS
1089	2439	A	8991	60	329	MALTPESPSSFPGLAATGSSVPEPPGGPNATL NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
		1				GVEDNAYTLEVNSRYMRAVGIM*IHL
1090	2440	A	8996	2	351	SNITITLT*MKKYDNTFCW*GCGQIG/T/LIYC
		,				WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL
ł						GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR
1091	2441	A	8997	97	456	LLIAALFIIVQYWKQSKDHYI YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT
1031	2441	Α	0277	31	450	LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG
						AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS
						FDKNYKAPIGADFEMERFEVLGIPF
1092	2442	A	8999	548	811	SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ
					:	FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG
1093	2443	A	9002	3	2745	ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS
						TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD
·				:		AMCCLRYWYTPESWICGGQWREYFSALRDF
						VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR
ĺ						LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLO
						HNCLGILGEDTDYLIYDTCPYFSISELCLESLD
J	]	}				TVMLCREKLCESLGLCVADLPLLACLLGNDII
						PEGMFESFRYKCLSSYTSVKENFDKKGNIILA
1						VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL
						*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP
}						GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG
						PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP
1		ľ				MCTGPEARQEVPMCTDSEP
1						RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY
L		L		L	L	TGPESRQEVLIRTDPESRQEIMCTGHESKQEV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	uence	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dellee		l	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	•	/=possible nucleotide deletion, \=possible
		i	L	sequence		nucleotide insertion
						PICTOPISKQEDSMCTHAEINQKLPVATDFEFK
		l		ļ		LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS DTEILKVARTHHVOAESYLVYNIMSSGEIECS
						NTLEDELDQALPSQAFIYRPIRQRVYSLLLED
						CODVTSTCLAVKEWFVYPGNPLRHPDLVRPL
		1		ĺ	1	QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
				1		CFNLSSSREELQAVESPFQALCCLLIYLFVQV
	j	Į	ļ	1		DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP
		l			1	DYINPRAVQLGSLLVRGLTTLVLVNSACGFP
						WKTSDFMPWNVFDGKLFHQKYLQSEKGYA
		1	٠		ĺ	VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1094	2444	A	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE
1074	2-7-7	1	7021	1	""	T*H\MAEPVSPLKHFVLAKKAITAIFDQLLEFV
					ļ	TEGSHFVEATYKNPELDRIATEDDLVEMQGY
						KDKLSIIGEVLSRRHMKVAFFGRTSSGKSSVI
	ĺ	1				NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA
						YLMTEGSDEKKSVKTVNQLAHALHMDKDLK
		1				AGCLVRVFWPKAKCALLRDDLVLVDGPGTD VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	A	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW
1095	2443	^	3022	1 *	557	LLHRRARRSSALCPRPRSWGVSGGEGAGARE
						P*ITSSSCCLSAA/SHLSIQSPNMAGARRRIRPQ
		1	Ì	1		LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF
		1	1	1		DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV
		<u> </u>				FAYGQT\GAGKTYTMGTGFD
1096	2446	Α	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL GOHSETPSLKKKVLAGYSGMCL*SQVLRRLRQ
						EDCLSPGGNCRES*SCPYTPAWITERDPV
1097	2447	A	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI
1		1			}	LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP
		l		ļ		GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG
						LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	Α	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS
			1			TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ
	Ì	l	ł	l		G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA
		ĺ				TMGEKRKVYYQKICYQKK
1099	2449	A	9043	185	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII
[					1	HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP
	<u> </u>				<u> </u>	RKIKTCPQNSCTSMLINAIHNDQKWKKINI
1100	2450	Α	9045	763	584	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL
110:	2451	-	1 0050	275	<del> </del>	SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	A	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS
٠		l	1			DSPASAS*VAGITDMCRYTOLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM
	1	1		1	1	DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF
	]	1			1	IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM
J	]	}	ļ		1	NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF
1			1	1	1	SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG
1			1		!	TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL*
]			1		1	LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS
}		}	1	1	1 .	SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
		Ī .	]	ļ	1	WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA
!	1	1	1		[	AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI
	1	1				ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  APFVLAVNC  KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
1105	2455	A	9065	366	778	YMQIGTTKNSRAT  DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR AWPCCPGWSAAWLTIVILAHYRRPGLERSCC LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL VLNS*TQGI
1106	2456	A	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT* AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	A	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA SAFPPAERSRGHRRASL*RARWSAAVPRRSA GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG QRPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL GAVAHSCNPSTLVGRGGRITRGOELR
1110	2460	Α	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS LLRKQRNKRMAIP
1111	2461	A	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP AAAGDPASLDFAQCLGYYGYSKFGNNNNYM NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK RAAPDSGKKPKTPKK
1113	2463	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE WRLQHKGRGRGDLHLPDHHLSVPSSADHPA QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPAPAEGGGTP*GSA GAGGP*GSPAGRACGAAGCRPRPPRPAASSA *NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTITELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP
	ļ				!	LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*
1118	2468	A	9154	471	2	GILTKLSAFLTIPRLQPHLIAALSPSS  AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	2	3187	ACPRLARRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQG\LRG\ WKARRA\TTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEEDEDDEDGGEEAPAPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1120	2470	Α	9163	124	207	PPRACRPCPRACPCPPT*KCSQPVSWPC
1121	2471	Α	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK V/CSHITDSLKFIGKGWVGMVTHACNPGTLG
1122	2472	С	9170	442	236	G*GGWIA*VREFETSLGNM MNRRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		j	914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
			j	residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
					[	WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
		İ				TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
	0.457.5	<u> </u>				IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
			ì			CSKPPKETGELENAESGGDGGRRGGKQDNV
						AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA LPMGFFYLYFRDPGREITWKHFVQYYLARGL
		1		1		VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE
						YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	Α	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
		l	İ	{		EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
		İ		1		ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
				<u> </u>		LTFFSTIS
1128	2478	Α	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
		ì				RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
			ļ			LILDGVPVALKKVQIFDLMDAKARADCIKEID
		]				LLKQLNHPNVIKYYASFIEDNELNIVLELADA GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
						ALEHMHSRRVMHRDIKPANVFITATGVVKLG
						DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
		1	ĺ			NG
1129	2479	Α	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
						PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
		Ì				RATT\KIRVVATITRARIEDMRHSATALTRPD
1100	2400	<u> </u>	1			ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
		]	1			LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
			:			DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
				100	000	PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
						CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
						GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
		ــــــــــــــــــــــــــــــــــــــ				AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK
		1			•	TQPVEATDDAFWDQFWADTATSVQDVFALV
		İ				PAAEIRAVREESPSNLATLCYKAVEKLVQGA
		J	]			ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD WRGFFWSTVPGAGRGGQGEEDDEHARPLAE
,		1				SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
						SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME
						LLKLLLTCFSEAMYLPPAPESWOH/RTHWFSS
		1				FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
						NHLY
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
			}			AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
						HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS
1134	2484	A	9210	66	1506	NVYFIV
1154	2404	A	9210	00	1586	MAGAGPKRALSAPVAEEKEEAREKIMAAK
			<u> </u>	!		RADGAAPAGEGEGVTLQGNITLLKGVAVIVV AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
						GVFSIVGALCYAELGTTISKSGGDYAYMLDV
		1	]			YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
		}				LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
				ř		YSVKAATRVQDAFAAAKLLALALIILLGFVOI
	1					
						GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID   SEQ ID   No. of nucle of nucle peptide pulse of peptide contide peptide contide control of nucleotide lust of peptide control of nucleotide lust of peptide pulse of peptide control of nucleotide lust of peptide peptide sequence   Peptide peptide   Peptide peptide   Peptide peptide   Peptide   Peptide peptide   Pep
eedide sequence USSN play56 opensed in get of first amino acid residue of peptide sequence
uence uence
Unence
International content of peptide residue of peptide residue of peptide sequence   TeThronine, V=Valine, W=Tryptophan, ide sequence   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   TeThronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Valine, W=Tryptophan, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=Deptide   Tethronine, V=D
residue of peptide sequence
Peptide   Sequence   Peptide   Pep
IVITLYYLTINLAYFITLSTEQMISSEAVAW    GNYHLGVMSWIPVFUGLSCSWOSLETS     RLFFVGSREGHLFSLSMIHPOLLTPVPSLVF     CVMTLF1AFSKDIFSVNFSFFNWLCVALL     GMWVRHKKPELERFRKVNLALPVFFLLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFFLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFTLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFTLACL     LLAVSFWKTTPWSVASDFTILLSGLPVYFTLACL     WWW.KKKWAPFGCHSSOPPOTLSKK     MOETLRILLSGLEVVETPENHCOETSSOPPOTLSKK     MOETLRILLSGLEVVETPENHCOETSSOPPOTLSKK     MOETLRILLSGLEVVETPENHCOETSSOPPOTLSKK     MOETLRILLSGLEVVETPENHCOETSSOPPOTLSKK     WYPIFFGTHET AFLGPKDLFPYDKCKDKYC     PKKKGPNEGL WEIQNNPHASYSAPPPVSS     SEAFANPADGSBADEDDEGMGVWAPTA     ATAASDRMESDSDSDKSSDINGGLKRKTPTA     A 9229 21 239 LFPRLECRPVTVNCTLNLPGSKNAPTTASC     GLYMNYNGGLASDGADFEPVSSSEEENSESS     SEKTSDDDTPEKKAAVKAPRRGPLGGRK     A 9239 21 239 LFPRLECRPVTVNCTLNLPGSKNAPTTASC     GLYMNYNGGLASDGADFEPVKTGPCKGPGGFKN     A 9229 21 239 LFPRLECRPVTVNCTLNLPGSKNAPTTASC     GLYMNYNGGLASDGAMFPPVKTGPCGGPGGFKN     A 9229 21 239 LFPRLECRPVTVNCTLNLPGSKNAPTTASC     GLYMNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     GRYWNYNGGLAV     A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN     GRYWNYNGGLAV     GRYWNYNG
GNYHLOYMSWIPYVGLSCFGSVNOSLFTS   RIPFYGSREGHLPSILSMIPPQLITYVPSULY
CVMTLPYARSKDIFSVINFRSFRINWLCVALA   GMWLBRRYBELERRIK VNLALPVFFILACI.   LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV.   LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV   WWKNKPKWAPPGHLSPRSPSCVRSSCMVP   WWKNKPKWAPPGHLSPRSPSCVRSSCMVP   DLVAFEDVAVNFTQEWSLLDPSQKNLYRF   MQETLRNLASIGEK WKDDIDQYKNPRN   RSLLGERVDENTEENHCGETSSQIPDDTLNK   RRFRSRYRGCSRFPRFGFLAVSMPHAFKP   DLVAFAKMGVPHWPABLDDIADGAVKPPF   KYPIFFGTHETAFLGPKDLPFYDKCKDKYC   PNKRKGFISGL WEIQNNPHASYSAPPVSS   SRAPANPADGSDADEDDEGIRGVMAVTAA   ATAASDRMESDSDSKSSDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SEKTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SSKSTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SSKSTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SSKSTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SSKSTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDNSGLKRKTPAL   MSVSKRARKASSDLDQASVSPSEEENSESS   SSKSTSDQPFTPEKKAAVRAPRROPICJGRK   APSASDSDSKASDSADGAKPPVAMARSASS   SSSSSDDVVKKPPRGRKPAEKPI-PKPGR   PKPEPPSSSSD   SKRTSDQPTTEKAAVRAPRROPICJGRK   ARSASAAPASCASSTISSRGQNKNTYGYGTLL   CTKHDWAPPRFCESSTTSDLEMRGGGRGKI   ARSAAAPASCASSTISSRGQLVKNNGGAN   GRRGSLNASGRTPPNCAAEDIKASPSSTNK   KNRPPMELDLINSSSEDNKPGKNRKGLNKRGAN   GRRGSLNASGRTPPNCAAEDIKASPSSTNK   KNRPPMELDLINSSSEDNKPGKNRKGLNKRGAN   GRRGSLNASGRTPPNCAAEDIKASPSSTNK   KNRPPMELDLINSSSEDNKPGKNRKGLNTAALDSCS   DISCEEGISNVALECSSPSTSVSAYDQLKAP   SPGAGNPFGTFKGRKEMSNOFGSIGAKA   NSGKKEGINNELNNLPVISNTAALDSCS   DISCEEGISNVALECSSPSTSVSAYDQLKAP   SPGAGNPFGTFKGRKEMSNOFGSIGAKA   NSGKKEGINNELNNLPVISNTAALDSCS   DISCEEGISNVALECSSPSTSVSAYDQLKAP   SPGAGNPFGTFKGRKENSPNOFARVATIVALAD   SPGAGNPFGTFKGRKGSPKMDAKTAALDSCS   DIFFIELD   GRESSFORM   SKALACKTTTITOTOLPSITITVQATPKSTT   TTPGGKPETTITOTOLPSITITVQATPKSTT   TTPGGKPETTITOTOLPSITITVQATPKSTT   TTPGGKPETTITOTOLPSITITVQATPKSTT   TTRAFGGFWRRAAAAGILPGKEAACLPSC/ VTAAVSGLLVGYEGIISGALLQIKTLLALS   HEQEMOVSSVJVGALU   STATATILITLIK   SKALACKTTTITAA   SPGASSALFGETLITGLIS
GMMURLRIRKPELERPIK VNLALPVEFILACL
LIAVSFWKTTPWSVASDFTIILSGLPVYFFGG   WWKNKPKWAPPGHLSPRSCVRSSCMVVP   1135   2485   A   9216   40   410   RDRLPPAYFCRPVVCVVTALDVGSPESQEM
1135   2485   A   9216   40   410   RDRI,PPAYFCRPVCVVTALDVGSPESQEM
DLVAFEDVAVNFTQEEWSLLDFSQKNLYRE
MOETLRNIASIGEKWKDQNIEDQYKNPRNN
RSILGERVDENTEENHCGETSSQIPDTILME   1136
1136 2486 A 9223 3 983 RRRRSRYRRCSRPPRGPLAVSMPHAFKEY DLVFAKMKGYPHWPARIDDIADGAVKPPP KYPIFFGTHETAFLGPKDLFPYDKCKDKYC PNKRKGFNEGL WELVINNPHASYSAPPPVSS SEAPEANPADGSADADEDDEGRGVMAVTAN ATAASDRMESDSDKSSDNSGLKRKTPAL MSVSKRARKASSDLDQASVYSPSEEENSESSS SEKTSDQDFTPEKKAAVRAPPRGPLGGRKK APSASDSDSKADSDGAKPEPVAMARSASSS SSSSDDVSVKKPPRGKPAEKPLPKPRGR PKPERPPSSSSD  1137 2487 A 9229 21 239 LFPRLECRDPVTVNCTLNLPGSKNAPTTASC GSTWNYRGGLPHFNFVKTGPRCSQAGLR RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCGVVTPVECTLNLPGSKNAPTTASC GSTWNYRGGLPHFNFVKTGPCSGAGLR RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCGVVTPEPCLGPCEPGTSVN EGIVWHETEGVLVVNVTVWRNKTYVGTLL CTKHDWAPPRFCESFTSDLENRGGRGRGK ARSAAAAPGSEASPTESFGLQNKNRGGANG GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSEDNKPGRKVRTNSRST TTPGGKPETTFLDGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTFKGKRELMSNGPGSUGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCAS DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTTITGIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRM/ KREADKVYTFTIDHAPSSIGS  1139 2489 A 9234 207 443 TRRGGPWRRAAAAGILPGREAAACLPSC/ VTAAVGGLLVGVEGGISGALLQIKTLLALS/ HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGGTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVMILGSLVRRKALAPRLLIRLIER TLRGGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGFEETLAFQGDSWNGVPSRAGLGI
DILVFAKMKGYPHWARIDDIADGAVKPPP KYPIFFGTHETAFLGPKDLEPYDKCKDKYC PNKRKGFNEGLWEIQNNPHASYSAPPPVSS SEAPEANPADGSDADEDDGRGVMAVTAY ATAASDRMESDSDSDKSSDNSGLKKFTPAL MSVSKRARKASSDLDQASVSPSEEENSESS SEKTSDQDFTPEKKAAVRAPRAGPLGGRKK APSASDSDSKASDGAKPEPVAMARSASSS SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSSD  1137 2487 A 9229 21 239 LFPRLECRDPYTVNCTLNLPGSKNAPTTASC GSTWNYRGGLPHPTNFFVKTGFRCSQAGLR RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN EGIVWHETEEGVLVVNVTWRNKTYVGTLL CTKHDWAPPRFCSPTSDLEMRGGRGRGK ARSAAAAPGSEASFTESRGLQNKNRGGANG GRRGSLNASGRTFPNCAAEDIKASPSSTNK KNEPPMEDLINSSEDNKPGKRVTINSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLLDKKNLGDKEKGK ANNCKTDKNPSKLKSARPIAPAPAPTPPQL IPTAFTTTTTGIDLPSLTTIVVQATPKSP KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHTLKDHLINNEGLANGLSESQESTM IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGLIPGREAAACLPSC/ VTAAVSGLL/VGPLGIISGALLQIKTLLALS HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVRRKALAPRLLIRLIEX RCGRGSPALFGRGAGRGGRFDTKCLA ATWGRLPGPETLIFQGDSWNGVPSRAGLG
PINKRKGPNEGL WEIQNNPHASYSAPPPVSS   SEAPEANPADGSDADEDDEG'RGVMAVTAY     ATAASDRMESDSDSKSSDNSGLKRKTPAL     MSYSKRARKASSDLDQASVSPSEEENSESSS     SEKTSDQDFTPEKKAAVRAPRRGPLGGRKK     APSASDSDSKASDDGAKPEPVAMARSASSS     SEKSSDDVSVKKPPRGRKPAEKPLPKPRGR     PKPERPPSSSSD     1137   2487   A   9229   21   239   LFPRLECRDPVTVNCTLNLPGSKNAPTTASC     GSTWNYRGGLPHFTNFFVKTGFRCSQAGLK     RGSREPPAWA     RGSREPPAWA     1138   2488   A   9231   1664   2   TRSVGVNTCEVGVVTEPECLGPCEPGTSVN     EGIVWHETEEGVLVVNVTWRNKTYVGTLL     CTKHDWAPPRFCESFTSDLEMRGGRGRGK     ARSAAAAPGSEASFTESRGLQNKNRGGANG     GRRGSLNASGRRTPPNCAAEDIKAGPGRGRGK     ARSAAAAPGSEASFTESRGLQNKNRGGANG     GRRGSLNASGRRTPPNCAAEDIKAGPGSRGKK     ARSAAAAPGSEASFTESRGLQNKNRGGANG     GRRGSLNASGRRTPPNCAAEDIKAGPGSIGAKAG     NSGKKGLNNELNNLPVISNMTAALDSCSA     DISDCEEGLSNVALECSEPSTSVSAYDQLKAP     SPGAGNPFGTFKGKRELMSNGPGSIIGAKAG     NSGKKGLNNELNNLPVISNMTAALDSCSA     DGSLAAEMPKLEAEGLIDKKNLGDKEKGK     ANNCKTDKNPSKLKSARPIAPAPATPPQL     PTATFTTTTTTTTTTGTIPGLPSLTTTTVQATPSSP     KPIQPKPTIMGEPITVNPALVSLKDKKKKEK     KLKDKEGKETGSPKMDAKLGKLEDSKGAS     DLPGHFLKDHLNKNEGLANGLSESQESRM/IKAEADKVYTFTDNAPSPSIGS     1139   2489   A   9234   207   443   TRRGQPWRRAAAGLLPGCAACLPSC/     VTAAVSGLLVGYELGIISGALLQIKTLLALS     HEQEMGVSSLVIGALL     1140   2490   A   9238   248   328   MAQGNNYGQTSNGVADESPNMLVYRKV     TURGHLGGASGRNTTGSLGEPQWLVATG     RPGTSPALFSGRGAATGGRQGRFDTKCLA     ATWGRLPGFEETLPQQDSWNGVPSRAGLGI
SEAPEANPADGSDADEDGGRGVMAVTAY ATAASDRMESDSDSNSSDNSGLKRKTPAL MSVSKRARKASSDLDQASVSPSEEENSESSS SEKTSDQDFTPEKKAAVRAPRGPLGGRKK APSASDSDSKASDDGAKPEPVAMARSASSS SSSSSDSDVSVKPPRGRKPAEKPLPKPRGR PKPERPPSSSSD  1137 2487 A 9229 21 239 LFPRLECRDPVTVNCTLNLPGSKNAPTTASC GSTWNYRGGLPHFTNFFVKTGFRCSQAGLK RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGNTCEVGVVTEPECLGPCEPGTSVN EGIVWHETEEGVLVVNVTWRNKTYVGTLL CTKHDWAPPRFCESPTSDLEMRGGRGRGKI ARSAAAAPGSEASFTESRGLQNKNRGGAN GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVKTNSRST TTPPQKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIGAKAG NSGKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGK ANNCKTDKNPSKLKSARPLAPAPAPTPPQL IPTATFTTTTTGIPGLPSLTTTVVQATPKSP KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRM IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/. VTAAVSGLLVGVELGIISGALLQIKTLLALS/ HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV TLRGHGGASGRNTTGSLGEPQWLVATG RPGTSPALFSGRGAATGGRQGGFPTKCLA ATWGRLPGPETLPGQDSWNGVPSRAGLGI
ATAASDRMESDSDKSSDNSGLKRKTPAL MSVSKRARKASSDLDQASVSPSEEENSESSS SEKTSDQDTPTEKKAAVRAPRRGFLGGRKK APSASDSDKADSDGAKPEPVAMARSASSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGR GSTWNYRGGLPHFTNFFVKTGFRCSQAGLK RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN EGIVWHETEEGVLVVNTWRNSTYVGTLL CTKHDWAPPRFCESPSDLEMRGGRGRGKI ARSAAAAPGSSASFTESRGLQNKNRGGANC GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPFGTFKGKRELMSNGPGSIIGAKAC NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNNPSKLKSARPIAPAPAPTPPQL IPTATFTTTTGIPGLPSLTTTVVQATPKSP KPIQPKPTIMGEPITVNPALVSLKDKKKKKK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRM/ IKAEADKVYTTIDNAPSPIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/ VTAAVSGLLVGYELGIISGALLQIKTLLALS HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKKALAPRLLRLRS TLRGHGGASGRNNTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRPDTKCLA ATWGRLPGFETLPQDSWNGVPSRAGLGI
MSVSKRARKASSDLDQASVSPSEEENSESSS SEKTSDQDFTPEKKAVVRAPRRGPLGGRKK APSASDBSKADSDGAKPEPVAMARSASSS SEKTSDQDFTPEKKAVVRAPRRGPLGGRKK APSASDBSKADSDGAKPEPVAMARSASSS SSSSDSDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSSSD
SEKTSDQDFTPEKKAAVRAPRGPLGGRKK APSASDSDSKADSDGAKPEPVAMARSASSS SSSSSDSVVKKPPRGRAPAEKPLPKPRGR PKPERPPSSSSSD  1137 2487 A 9229 21 239 LFPRLECRDPYTVNCTLNLPGSKNAPTTASC GSTWNYKGLPHPTNFFVKTGFRCSQAGLK RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN EGIVWHETEEGVLVVNVTWRNKTYVGTLL CTKHDWAPPRFCESPTSDLEMRGGRGRGKI ARSAAAAPGSEASFTESRGLQNKNRGGAN GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMEIDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEGGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAC NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQL IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKOAS DLPGHFLKDHILNKNEGLANGLSESQESRM IKAEADKVYTFTDNAPSSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAGILPGREAAACLPSC/ VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYKKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKKALAPRLLLRLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRPDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
APSASDSDKADSDGAKPEPVAMARSASSS SSSSSDDVSVKKPPRGRKPAEKPLPKPRGR PKPERPPSSSSD  1137 2487 A 9229 21 239 LFPRLECRDPVTVNCTLNLPGSKNAPTTASC GSTWNYRGGLPHPTNFFVKTGFRCSQAGLK RGSREPPAWA  1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVN EGIVWHETEEGVLVVNVTWRNKTYVGTLL CTKHDWAPPRFCESPTSDLEMRGGRGRGKI ARSAAAAPGSEASFTESRGLQNKNRGGAN GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNOPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQL IPTATFTTTTGTIPGLPSLTTTVVQATPKSPI KAPADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/ VTAAVSGLLVGYELGIISGALLQIKTLLALS HEGEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRS TLRGHGGASGRNVTTGSLGEPQWURVATG RPGTSPALFSGRGAATGGRQGGRPDTKCLA
SSSSSDDVSVKKPPRGRKPAEKPLPKPRGR
PKPERPPSSSSD
GSTWNYRGGLPHPTNFFVKTGFRCSQAGLERGSREPPAWA
RGSREPPAWA
1138
EGIVWHETEEGVLVVNVTWRNKTYVGTLL CTKHDWAPPRFCESPTSDLEMRGGRGKI ARSAAAAPGSEASFTESRGLQNKNRGGANG GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSLIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNYPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRM/ IKAEADKVYTFIDNAPSSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC// VTAAVSGLLVGYELGIISGALLQIKTLLALS/ HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
CTKHDWAPPRFCESPTSDLEMRGGRGRGKI ARSAAAAPGSEASFTESRGLQNKNRGGANG GRRGSLNASGRTPPNCAAEDIKASPSTINK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPFGTPKGKRELMSNGPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMAI IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRLRLSG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
ARSAAAAPGSEASFTESRGLQNKNRGGANG GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAG NSGKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTAIFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC// VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLIRLILSG RPGTSPALFSGRGAATGGRQGGRPDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
GRRGSLNASGRRTPPNCAAEDIKASPSSTNK KNKPPMELDLNSSSEDNKPGKRVRTNSRST TTPQGKPETTIFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRM/ IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/ VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
KNKPPMELDLNSSSEDNKPGKRVRTNSRST. TTPQGKPETTFLDQGCSSPVLIDCPHPNCNK YKHINGLRYHQAHAHL.DPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEKK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKKALAPRLLIRLLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
YKHINGLRYHQAHAHLDPENKLEFEPDSED ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC// VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
ISDCEEGLSNVALECSEPSTSVSAYDQLKAP SPGAGNPPGTPKGKRELMSNGPGSIIGAKAG NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
SPGAGNPPGTPKGKRELMSNGPGSIIĞAKAC NSGKKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSC TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
NSGKKGLNNELNNLPVISNMTAALDSCSA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLRSC TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
DGSLAAEMPKLEAEGLIDKKNLGDKEKGKI ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  TRRGQPWRRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  I140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  I141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
ANNCKTDKNPSKLKSARPIAPAPAPTPPQLI IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRLRLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPI KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
KPIQPKPTIMGEPITVNPALVSLKDKKKKEK KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
KLKDKEGKETGSPKMDAKLGKLEDSKGAS DLPGHFLKDHLNKNEGLANGLSESQESRMA IKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCA VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
DLPGHFLKDHLNKNEGLANGLSESQESRMAIKAEADKVYTFTDNAPSPSIGS  1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSCAVTAAVSGLLVGYELGIISGALLQIKTLLALSGHEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSGRAGASGRNVTTGSLGEPQWLRVATGRPGTSPALFSGRGAATGGRQGGRFDTKCLAATWGRLPGPEETLPGQDSWNGVPSRAGLGI
1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC// VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
VTAAVSGLLVGYELGIISGALLQIKTLLALSG HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSG TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
HEQEMGVSSLVIGALL  1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV  1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLGI
1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRS TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLG
TLRGHGGASGRNVTTGSLGEPQWLRVATG RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLG
RPGTSPALFSGRGAATGGRQGGRFDTKCLA ATWGRLPGPEETLPGQDSWNGVPSRAGLG
ATWGRLPGPEETLPGQDSWNGVPSRAGLG
WPWAAALVVHCYSKSPSNKDAALLEAARA
WI WAAAL VIIIC ISKOI SINKDAALLEAANU WMQEVSRNRCALLHSAAVQEYGYGN
1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGC
FALGNDFLNITTKAQA/TKEKLDKLDFIKIKT
CTSMDAIEKTEPLTKWTKAFVSHVSYKRLL
GICKEYSRQ
1143 2493 A 9247 264 115 GLPQQTSTIQPPGAPI ARDSTSIIRMGPEIPPP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE AMEESDRPCEISEIDDNPKISENPRRSPTHEKN TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM ERRR
1145	2495	A	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL WDTAGQERFISIT
1146	2496	A	9277	592	814	MFTYLEGREGIKSOPKMEPHSVTVRLECSGMI SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEEKEEEEEEGWVNGMENSHPP HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ HLHSTSVMGNIIHVELDTKGETRMRFYELLV TGRYTPQTLPVGELDAVSPIVNETLQLSDALK RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY NRRHEHHYVHNSPAVTAVAGATAAFRGSSD LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN LSLPDVGFTSTTVPK\MIVDI\QSRSRVISYAG CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF GGMEERHAPECDGL
1149	2499	A	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE FRLVAADRSMGRYMLFGVINLICTGFLLMWC SSTNSIALTISYTYLTIFDLFSLMTCLISYWVTL RKPSPVYSFGFERLEVLAVFASTVLAQLGALF ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF TMLSIRNKPFAYVSEAASTSWLQEHVADLSR SLCGIPGLSSIFLPRMNPFVLIDLAGAFALCIT YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL
1151	2501	A	9309	205	406	SLP VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR PPPHR\(\text{EIFFVFLAETGFHRASQAGPDLPTS/S/I}\) PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL GTNVSLRAA
1153	2503	A	9315	392	I	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR PGNS

<u> </u>	LOTO TO	134-4	LOCO	Predicted	D 3' 4 3 3	1 4 4 4 1 1 0 0 0 A
SEQ ID	SEQ ID	Met hod	SEQ ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	поа	1 -	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide		in USSN	location		F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq-	seq- uence	i	09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Derice		1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ		]	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	Soquenoo	/=possible nucleotide deletion, \=possible
		ļ	İ	sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
1			1	1	, , , ,	PT
1155	2505	A	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	A	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
						EAAQRHCRASVSILRMRRPGQGSSRPARVPL
	1	Í	1			RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGSHPAGAQPGGRAIGAGWQS
		ļ			1	KEPLWEGLORSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
						LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
						VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
			I	1		SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
ľ		1	1	İ	ĺ	RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNOVNGNGNOLKRKHESMCPVSLTONTVR
		ļ		Ì		LMEAGLPQKQAERADELFEAGLVIYVKLDER
	}	l	}		j	VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
ł		1	İ			KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
1			1		,	DHTDQELREEIHKANVERVVHDVSQEATIEKI
1		ļ	ł		ł	RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
			1			EAELPIMSQLTEIETCVEC
1161	2511	A	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
	l					AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
					·	DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
	j				1	QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
}		ļ	ł	1		GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
1			1			TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
						VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
İ		ĺ	ĺ	İ	ĺ	VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
						GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
			l .		ļ	FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
ļ		)		Į		WIDGPKDTSVEDALERTYVPCLKTLQEEVME
				•		AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
						G
1163	2513	A	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
		l	ł	ì		LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
		•	1			AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
	2.00	<u> </u>	<del> </del>			FSSYQRNNPDLILNDTIMPNIK
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
1		ſ .	1		[	HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
		ļ.	1		1	FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
		1	1	<u> </u>		VAGASVGAGVWARNPRYRTEGEACVEFKA
1		Į .	1	]	1	MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
	1			ŀ		VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
			1			SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
	1	1	1	[	]	SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
1		l	1	1		VTMAISWITTVVPLLTFEVLLVHRLDGHNTFS
1				ŀ	1	YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
			Į.			AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
1103	2313	^	3302	J41	ופכן	VPEGVRLADGPGHCKGRVEVKHQNQWYTV
1		1	1	1		CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
		1				TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
1.	1	1	1	1	1	LGEDTLFHVEYTSVHGRERLSAKD
1166	2516		9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
****	2310	^	1333	201	1 307	GAISAHLAHCNLCLPGSSDSPASAFQVAS
1167	2517	A	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP
	1 201/	<u>^^</u> _	1,500	1 101	1007	DIT ODO! OO! DICH RE DRI IFURRDEDHIFF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		}		peptide		/=possible nucleotide deletion, \=possible
			ļ	sequence		nucleotide insertion
				i		PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
	ļ	l			1	ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
						PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS
		}				PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID
						LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS
		l		}	1	QQSILAGLVVVATTGMIGSPLECLFGELGGRA
	1					DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL
		İ				FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN
ł		}				NSARRMEAMASGSNWLSGVNVVLVMAYWS
	1	Ì				LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH
		<u> </u>	1	]		NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170	2520	Α	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR
	1	1	1			ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF
						TNETWQARTGEPLPDHLVLLMWSLIVSLYPL
ŀ		1			ĺ	GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS
	ł	}			ļ	AAILFGFSRKAGSFEMIMLGRLASWGVNAGV
					ļ	SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA
ļ		1				LGIVMGQVVGLSTTAATGLRGL\AGELEELEE
		}				ERAACQGCRARRPWELFQHRALRRQVTSLV
						VLGSAMELCGNDSVYAYASSVFRKAGVPEA
	Į	1				KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL
	0.504	ļ.,				WGGTPRSFALNQFTLQKKKK
1171	2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA
		ļ			1	LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE
						TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV
					}	EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG
1172	2522	A	9384	20	266	SKRLFFWMQEP
11/2	2322	^	9304	20	355	GWNGRSTEASPAAEAPHVPHKET\KAAMGTQ
			1		}	CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL
						LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL ASFNEVGNTALIVLESY
1173	2523	A	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ
11/3	2323	Α .	2323	450	87	QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI*
	1	ĺ				KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV
			Ì			
1174	2524	A	9397	77	374	IRPPISFSKINNGP   ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ
		٠. ا		<i>''</i>	3,7	RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF
		İ		1		ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF
		1				LGML
1175	2525	A	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT
· <b>-</b>		l	1	]	<i></i>	AKMNKWSKGKVRDKLNNLVLFDTATYDKL
				,		CKEVPNYKLITLAVVSERLKIPGSLARAALHE
ł	ł	ł				LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY
				-	~	GERGYAQNGDF*DAQLDDYSFSCYSHAQVN
	ł	}				GAPNSLTRAYDDP*VKISGLECQKVGALVEV
	<b>[</b>					KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV
		}		-		YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD
						FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF
l	1		1	1		DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1170	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP
1178	ı -					ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL
11/6	l	1		,		
11/6						ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAFI
1178	2529	A	9420	1450	1655	ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL LSSAGTKMNLN*KNYWPGASAHACNPSTLG
	2529	A	9420	1450	1655	ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide					to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	Ī	09/496	correspondi	1	
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	ł	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i		1	peptide	•	/=possible nucleotide deletion, \=possible
ľ	ľ	ľ	ľ	sequence	ĺ	nucleotide insertion
L		ļ., —	0400		200	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
1180	2530	A	9422	176	375	
l			j			SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
i		ļ				L
1181	2531	A	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
	====	**		-		IARTILDRLTGIPHGYCFVE*ADWATADKCVH
ļ	ł	ŀ	į.		i e	IYNGKPLPGATPLLSLQLHQLAHLGS
		L		L		TINGKFLFGATFLLSEQUIQUATEGS
1182	2532	Α	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
1			į.			SMILK*MGAGDEKISAMGKARVDHRELYLGL
•	İ	İ				LYPTEDYKLTFRARH
1183	2533	A	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR
1103	2333	^		707	-	KGCSGWAPWLSLQCQHFGRPRWADHLRSGV
1			1			
	1	1	1			RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
		{			1	ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
		İ		1		ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG
1104	2334	^	7702	1 331	1 333	RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
[	1		1			
	l		<u> </u>	l		IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
i		ı	1	ļ		WWWGWECWVRALKLSSGPAGPLACWVAK
ł		1			1	KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
l	İ	1				WSAVVQFWLTAASNSCFSLLSSWDYRCA
		<b>.</b>			L	
1186	2536	Α	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR
	Ì		1			GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	A	9469	388	3	EVAPGPSQILPRRVTLGGDRPQFSLPGPRLPQ
1107	1 200.	1 **			*	SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
ł		ł	1			NPASPHPEAPQEPWDSASGSVGSFSLGRGAK
ł		İ	1	Ì	į.	ACCUMON CONCORDO CONTENTO DE LA CANTANTA
1	j					ASS*VPGKGRGPRQGSELLAETILELFLALAN
ł	1	ł	ł		1	S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
		1	1			GRLMANPEALKILSAITQPMVEEAIAGLYRAC
l	Į.		1			*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1100	0530	-	9480	604	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
1189	2539	Α	9480	584	עסי	
			i			PSLLKIQKISWAWWRAPVVPATWEAEAEEW
ł	ł	1	ł	1		R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
1	1 22	1		100	**	PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR
	i	1	1	i		GASSCRRRCNPVLAARKAGSPRSHSTRENC
1	1	1		1		
<u></u>	<b>.</b>		L	<u> </u>	<u></u>	RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	Α	9489	1	411	LADALCLSAAATGAVRPGARAQPSTRRRLSP
	1	1		I		SVRVCCRAAAASNLLYSSCLQRHSERASEEG
	1	1 '	1	I	1	ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
		1	1			MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ
1		1	1	1		
L		L	L		ļ	KEEELTAVNVK
1192	2542	A	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
1	1	[	1	]		CEEDERKMAREFLAEFMSTYVMMNIHMIVE
1	ì	1	1			KDTYSDHEEINTS
1100	0545	<del> </del>	10500	106	<del>                                     </del>	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
1193	2543	A	9509	186	1	
			J		l	FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	Α	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL
1		1	1		ļ	LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
1	ł	1	1	1		SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
1	l	1	1		i	
			L	<u></u>	<u> </u>	RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
1		1	1			AGGGASLPVÄAGSCAAAPHTEPGAPOHLLDC
1	1	l	ŀ	l .		PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
ļ	1	1	1	1		LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
1	1	1	1	]	l	
		1	1		ì	PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP
		1	1			APGSGPCGATARPSRGGRAGGSRARRPIPPGP
1		1	1			GTRRTPSGCQNPAASGG
				<del></del>		<u> </u>

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ		peptide		/=possible nucleotide deletion, \=possible
110	2516			sequence		nucleotide insertion
1196	2546	A	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
	1	ļ	ļ		i	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
1197	2547	A	9521	289	448	GYIGALFPMSGGWPGGQ IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
1,	2547	^	7521	207	740	HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
1		1				APNWKYKYGY*IPVDMLC
1198	2548	A	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
					] -	VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
)	}		]		].	SSYS
1199	2549	A	9546	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
						V*QRGDGKNPGVTHLNRPVGTX
1200	2550	Α	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
						KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	2551	A	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
1						GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
	l		\			YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
ľ						KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
	1					PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK PDVHFFHCDEVEAELVHEYMESALTDCRLGK
	Ì					AMRP
1202	2552	A	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
			"	.20	*	LDCERPPQGPLPSLPELAKTSYSDLTGLATED
1						*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
ł	·		<b>!</b>			LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
						SKPRATPPLFCSLHTF
1203	2553	Α	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
						EKEKRRREKGEREERKMRHRERKGESGQRD
1001						TMENWRVERLTEKER
1204	2554	Α	9573	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
	1					DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
						EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
1205	2555	Α	9577	64	424	HDCRHKEDAGVICSEFTALR ARGSCPTRPRTANGRMGETKDAPQMLVTFK
1.205	2333	. 22	/3//	04	727	DVAVTFFREEWRQLVLVHRTLYR*GMLETC
						GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
						SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
Į.			]			NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
				İ		GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
1207	2557		0506		410	YNPGLPPLRTWNGQKLLWL
1207	2557	Α	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
1						PDGCRNVLRPKYYRLCDKAESWGIALETVPT
						GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
ļ				j		FGILFSICFS
1208	2558	A	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
					-	FADAWADAW
1209	2559	Α	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
l				i		GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
1						RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
				J		MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
1010	0500		0672	204		LLNASITETFNC
1210	2560	Α	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
1						DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
						KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS
				ļ		GELSAK
1211	2561	A	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR
			للتتتا			AAYSAASSAAMII AISAATA IMIGAASSAASAAA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST EENFGEKLHDIGFGNGFLDKT*KAQATKAKI DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI NHLPETERNLLEHGLMYIRLNAAFCSLVAHS LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV EEHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL GALGGRGGRALGGSRWPP°LPGETLFSGCKH RRRRGSDAAPGEEAGT
1218	2568	A .	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL
1220	2570	A	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K
1221	2571	A	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS
1222	2572	A	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE
1224	2574	A	9700	3	632	DAWASGELGSLFDHHVQRAVCDTRAKYRE GRRPRAVKVYTINLESQYLLIQGVPAVGVMK ELVERFALYGAIEQYNALDEYPAEDFTEVYLI KFMNLQSARTAKRKMDEQSFFGGLLHVCYA PEFETVEETRKKLQMRKAYVVKTTENKDHY VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
1005	0.555	<u> </u>	0510	sequence		nucleotide insertion
1225	2575	Α	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
1226	2576		0712	-00	100	TDDVHINDISLQDYIAGVRLILL
1220	2376	Α	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT
'		l				ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
						ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
1		,				WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
1007	0.500		0500			MAFVFSSLI
1227	2577	Α	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
						LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
	•	}				NRKKMKDCQLRKQQNENVSRAVCALLNSGG
						GVIKAEVENKGYSYKKDGIGLDLENSFSNML
1000	0.570		0000	0.50		PFVPNFLDFMQNGNYF
1228	2578	Α	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
		<u> </u>				TLVLQKSDVEAVF
1229	2579	Α	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
			1			GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
		ľ				QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
			1	-		PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
]						NETDLAGPMVFCLAFGATLLLAGKIQFGYVY
						GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
						GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
						WCSFSASKIFISALAMEGQQLLVAYPCALLYG
1000	2.222					VFALISVF
1230	2580	A	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
						HFSPERPFMDYFDGVLMFVDISGKCKRDVCL
1,221	0.501		0744			MWMSNRLAWEFTCRA
1231	2581	Α	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
						ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
						LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
1 1				•		LRCGWSPAEELNYTVPGPGPAGEASPRQCRR
1						YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
1						RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ
						SSVNVGFFIGSMSIGYLADRFGRKLCLLTTVLI
						NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
						WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL
1						LVLAGVAYALPHWRWLQFTVALPNFFFLLY
						YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG
1232	2582	Ā	9753	164	517	RSLPASL PROMOCOPPOTENTS WELL PROVIDED AND CO.
1.22	2002	Α.	7133	101	217	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
[ [	i					YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
						LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS
1233	2583	A	9757	25	419	FGILLWSGAGLSSSFISPRYSWLF
,	2002	^	7131	ا	417	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
					l	IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
]	-					KRAYKSYVRALPLLKKMGINSILLRKSIGALE
						VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
1234	2584	A	9765	71	456	PLEI DWGESI HEI BWAWI CDI SSCRENORION
14.24	2004	^	7/03	/1	430	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
j						CSRCGYGVYPAEKISCIDQIWHKACFHCEVC
	- 1				ĺ	KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
						VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
1235	2585	Α	9767	52	559	FHWD
	2000	^	2101	ےد	229	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
						SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
						ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
1	1			ł	ŀ	CARPGHSARNNTDKNLPHTAILLVTSNTYTTI
			1	l		KINFQAGRSGSCL
1236	2586	A	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC
	2200		2110	J. J. L.	000	LUGEALI VALLIAATIGE I ASNIESI YKKHLC

				T		<del></del>
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		] .	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		•	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l .	l	1			Sequence	
1	Į.		1	peptide	1	/=possible nucleotide deletion, \=possible
<u> </u>	<u> </u>	<u> </u>	<u> </u>	sequence		nucleotide insertion
	1					LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
1	1				ļ	DGFVIVYDISDRSSFAFAKALI
1237	2587	Α	9793	266	515	NILAIIYFPFPRLFLLRDSQSNPKAFALTLCHH
	ļ	ł				QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
		i				RYKPTRPVCITQFQGCS
1238	2588	A	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
1236	2300	Α .	3602	337	307	
	ļ				1	FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT
1		ł				TDEDSTRFQIINEASKVPLLAEIYGIEGNIFRLK
	1	İ			i	INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
	1				į	GSLILADGKGDLKC
1239	2589	A	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
İ	1					MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
ŀ					1	GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
	ŀ	1				PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
ļ		i	ĺ	1		DIKFLTKGDNNEGDDRGSYK
1240	2590	<del> </del>	9819		305	
1240	2590	Α	9819	3	303	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
		l			1	PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
	1	1				YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
,	ļ	)			J	DPNSEEHG
1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
					i	LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
1		1				GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
ł		ł	1			SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3	589	<u> </u>
1242	2392	A	9843	3	289	TISCGPATEPPASLLSSASSDDFCI:EKTEDRYS
i		1				LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
	1.	l				DGSVKCTHQPCLIKWISERGCWSCELCYYKY
}	1	1	1		]	HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
1	1	l		ľ		LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
1	i	1				MYGFMDVMIVAVDSEDMVQAAKEVGKRWS
i	İ			Į.		DIPP
1243	2593	Α	9846	198	411	WRISHHAGKMPVMKGLLAPQNTFLDTIATRF
		1				DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
		Į	1	l		FARTEVMQ
1244	2594	A	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
1244	2394	I A	7040	110	050	•
1	į	l	1	<b>}</b>		PANGAGGPGGASARKLSTFLGVVVPTVLSMF
1	I			İ		SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA
	1			t		LTVLSVCAIATNGAVQGGGAYCILQHRWTG
		1	1	l		VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
		ł	1	l	ł	NVCGCAVSLLGLVESVLDVFGA
1245	2595	A	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
1			1		1	AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
1	1		1	l		SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
	1	ŀ	Í	-	ĺ	AMSLTEDVLAAALADHLPEDK WSAEKRRPL
		1		1		== - = - = - = - = - = - = - = - = - =
		1	1			KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
			1	1		YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
		l	1	1		RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
1	1	ł	1	!	ł	SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
		Ī				NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
		ŀ		İ	1	EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
		1		1		LSGPTSEGLMTWELDRLLWARSVENLATATT
ľ		ĺ	-	ĺ	l	TLTSLA
1246	2506	-	0050	114	464	
1246	2596	Α	9850	114	464	PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS
}		l		1		ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
		l		[		QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
<u> </u>		}		<b>)</b>		RRERPAGADSLSWGAGPRISSYV
1247	2597	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
					1	HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
		1	1			GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
ł	1	l	1	1	}	LYKVFKYSSRCTS
					1	DIWINIONOIO

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	The said and the s
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	<b>I</b> .	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	dente		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uchee			314	amino acid	of peptide	Q=Glutamine, K=Arginine, S=Serine,
	Ì	1		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
	1	İ	Į.		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1	1	peptide		/=possible nucleotide deletion, \=possible
1248	2598	<u> </u>	0053	sequence		nucleotide insertion
1248	2398	Α	9853	58	444	RVDDFVYSKGGKDAGGADVSLACRRQSIPEE
		1		İ		FRGITVVELIKKEGSTLGLTISGGTDKDGKPR
			1			VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR
	1		1	1	!	LRHDEIITLLKNVGERVVLEVEYELPPPGGCP
	<del></del>	<u> </u>			1	WT .
1249	2599	A	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV
	į		ļ			RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA
	i	İ	1			SGSGVAAGPAARHAPRRRCADAGEAVGASC
	1	1	ŀ	1		GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT
	·	-				PMGAGDAGASAESAVTTAPQEPPARPLQAGS
		1	ļ	1	}	GAGPAPGRAMRSTTLLALLALVLLYLVSGAL
	,					VFRALEQPHEQQAQREL'GEVREKFLRAHPCV
					!	SDQELGLLIKEVADALGGGADPETNSTSNSSH
				1	1	SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK
	1	1	ļ	1		ELPHGGRCRETEGSQVAPRLPASPLCPGYGN
						VALRTDAGRLFCIFYALVGIPLFGILLAGVGD
		ł	ŀ	!		RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA
		1		[	ĺ	MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY
		l				FVIVTLTTVGFGDYVA
1250	2600	Α	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF
		1	70.5	-	032	EWYYTDQPHTQRRKEILAKYPAIKALMRPDP
		{			1	RLKWAVLVLVLVQMLACWLVRGLAWRWLL
						FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR
	1	ĺ				
						AARNRWLAVFANLPEGVPYAASFKKYHVDH
	İ	(			ĺ	HRYLGGDGLDVDVPTRLEGWFFCTPARKLL
		1				WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV
1251	2601	A	9875	150	1209	QLA
1231	2001	^	3673	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR
			i			LESYRPDTDLSREDTGCNLQHISDRENIDDLN
						MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN
	1	l				HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI
						KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL
						SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL
						TAECAIVTLVYLERLLTYAEIDICPANWKRIV
	ł	1				LGAILLASKVWDDQAVWNVDYCQILKDITVE
	1					DMNELERQFLELLQFNINVPSSVYAKYYFDL
		I				RSLAEANNLSFPLEPLSRERAHKLEAISRLCED
1055	0.00	<b> </b>	1			KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	Α	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT
	1	1				KAVGIKMGSLSTANVEFCLDVFKELNSNNIG
	1	Ì				DNIFFSSLSLLYALSMVLLGARGETEEQLEKV
	<u> </u>					WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC
		l				NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA
		]	j	-		AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE
	-	1			• •	STELSATTFSTQSPLQKLFARKMKILGTIQILF
	1	l				GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG
						SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA
	1	l		i i		LGAIAGULLTFEFHPRSKLHL
1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT
	2003	^`	7070	, 4	300	GATCVCI DAVGMCDOT SCAT TOT STATE OF STATE
		ŀ				GATCVGLPNVGMCPQLSGALTFMYLQQGNQ
	1	İ	1	l		EATVAPDTMAQPYASAQFAPPQNGIPGEYTA
1256	2606		0000	05		PHPHPAPEYTGQTT
1256	2606	Α	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG
	1			: [		DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK
			] 1	1		· · · · · · · · · · · · · · · · · · ·
						PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH
1257	2607	A	9905	374	459	PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	j	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN 09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	uence	Ì	914	correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Ì	1	}	1	residue of	sequence.	Y=Tyrosine, X=Unknown, *=Stop codon,
}			ļ	peptide		/=possible nucleotide deletion, \=possible
1			İ	sequence		nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH
						QRRGPSCGASGDPQCVGSPHPQRARPLLARP
1		i		ł	Į	GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS
						YLGECGSSSYVTGAACISPVLRCREWFEAGLP
Ì	l		l			WPYERGFLLHQKIALSRYATALEDTVDTSRL
		i				FRSRSLREFEEALFCHTKSFPISWDAYWDRND
		[	ĺ		ĺ	PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT
	}	]		j		ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS
1	-					HEVILESFRALTEFFRTEERIKGLSRHRASFLG
1		1	Ì	)		GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR
			}	ļ		YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR
		1		İ		ERPAARSGPEMRVRYPVVAAVLAPYLALSQD
	ļ	ļ		1		PMVKSSASGQGASGSYNHVREEMLIKAGGA
1		1				MSRRVVRQSKFRHVFGQAAKADQAYEDIRV
1		l	ŀ	ł		SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL
1000	2600	<del> </del>	0010	-	025	PLAK  CCEVEICESTCCWIEDSSVTTCCVVAV ADD AA
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL
	j	1		ł		PPRPGRSHRKRKLVSTK
1260	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR
		]		]	1	LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ
		1	1		[	IQEKVRYMSSTDAKVEQMFQILLDPEEYGTE
	ļ	ļ			Ì	KKQKASQNLVVLAREDAGAEKIFRSNGVQLL
	į	1	]	}		QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV
		ŀ				MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG
				-		RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
		1	ł	1	1	PTRVDHNGALLAFSPPPPQRQRRGTGATAES
						RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY
1260	2612	<u> </u>	0021	160	425	WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLOELOGRHALHTFSLERTCSYQDFL
		ĺ	1		[	WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG
1				-	, , ,	CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY
	<u></u>	L		L		YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG
				1		WLALLLGALLGTAWARRSQDLHCGACKAVR RRVROFNIYDY
1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT
1203	2013	^	7770	*	ئے۔ ا	EVAGEEELQMIQPEKLLLVTVGKTATLHCTV
1		}	1	1	1	TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP
	1					RVTTVSDLTKRNNMDFSIRISSITPADVGTYY
1	1	ľ		(	1	CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG
10.5		ļ. <u>.</u>	1.0000	L	L	FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG
120/	2017	^	10004	30	""	PGPGFGFASKTKKKHFVQQKVKVFRAADPLV
		[				GVFLWGVAHSINELSQVPPPVMLLPDDFKAS
	]	]	1	}		SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL
	1		1	1		RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS
		1		ļ		YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS
1260	10010	<del> </del>	10005	<u> </u>	1000	SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR
L	<u> </u>	ــــــــــــــــــــــــــــــــــــــ	<u></u>	<u> </u>	<u> </u>	SANTERISTORSWAGGWAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

	1 000 10		T == 0	<u> </u>		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	1	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иелсе	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
[	(	l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			Į	peptide		/=possible nucleotide deletion, \=possible
L				sequence	<b>{</b>	nucleotide insertion
						TKVSLGLA
1269	2619	Α	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
1						LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
	ł	ŀ	1			IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
	]		1			AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
	]				ļ	VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	DUDDEUDU DDCL MODOD MOVID COM MONIO
1270	2020	^	10011	2	200	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
	1	Ì	l			PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
ł		ŀ	i			SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
	1		l			RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
ĺ	1	ĺ				QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
1	l					QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
		<u></u>	L			DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
l						FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
1 .	}		]			LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR
						LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
1	ł	ł	}			SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	
12/3	2025	Λ.	10010	•	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
]	}		j j			SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
						AQAAREPAAFWGPLARDTLVWDTPYHTVW
	ł					DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
			1			PESVALIWERDEPGTEVRITYRELLETTCRLA
1	1		1			NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
						CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
	i		1 1			ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH
}	1		ł l			VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
l i			<b>!</b>			VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
1 1						QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
!						WITGHSYVVYGPLCNGATSVLFESTPVYPNA
						GRYWETVERLKINQFYGAPTAVRLLLKYGD
)			i 1			AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
						RVVGDSRCTLVDTWWQT
1274	2624	A	10017	1	3750	
1-/-	2024	A	10017	1	3730	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG
						KTLGSFFGSLPGFSSARNLVANAHSSARARPA
						ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
						KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
				ĺ		GVASVVDVAKGVVQGGLDTTRSALTGTKEV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG
				İ		TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
		1	l		j	LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
						KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
				ł	ł	DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
				ļ		GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
						IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
				}		KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
						LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
				ļ		
						AANVAKGAMQTGLNTTQNIATGTKDTVCSG
			1			VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
						SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
						AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
				ļ		TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
				İ		VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
	ľ		·	I	}	KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
			•	1		DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT
			1	1		GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
··· ]			1	· .	J	QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK
· I						GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN
	1					

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion  VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM
1275	2625	A	10025	124	415	NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL QDCFRL TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG
1276	2626	A	10030	3	507	KKPFS GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYDI. LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDQ
1277	2627	A	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL
1283	2633	A	10088	316	516	ALEGPLLITFFVRRVKGTFGEIM MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	A	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	A	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFOL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	Α	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1,000	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	deixe	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence		ĺ	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l	ĺ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		[			Sequence	/=possible nucleotide deletion, \=possible
1		ĺ	[	peptide	[	nucleotide insertion
	ļ	ļ	<u></u>	sequence		LGNVLTSTPNAKTVNGKAESSDSGAESEEEE
Į.				ļ·		
1000	2642	<del> </del> -	10104			AC
1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC
	ı		i i	1		KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL
}	ŀ		ł	1	}	CGRGIRDSARMCSTCACVEYYGKALEVLVK
i	i					GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV
ì		Ì			•	CTDVMARGIDIPEVNWVLQYDPPSNASAFVH
1	i	<u> </u>			ļ	RCGRTARIGHGGSALVFLLPMEESYINFLAIN
•		Ì				QKCPLQEMKPQRNTADLLPKLKSMALADRA
1	ļ	l		į	İ	VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
		ļ				DFASLARGFALLRMPKMPELRGKQFPDFVPV
		1		İ		DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN
1	·			1		EGRRKFIKNKAWSKQKAKKK
1294	2644	A	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA
	)	l				ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ
ı		1	ļ	ļ		LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF
1						GVLFALCFSCLLAHASNLVKLVRGCVSFSWT
1		Ì				TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN
1	{	<i>'</i>	1	ĺ		MTPCOLNVDFVVLLVYVLFLMALTFFVSKAT
				i	ļ	FCGPCENWKQHGRLIFITVLFSIIIWVVWISML
	]	!	Ì			LRGNPQFQRQPQWDDPVVCIALVTNAWVFL
	ł	Ì		1	i	LLYIVPELCILYRSCRQECPLQGNACPVTAYQ
	ļ	1	٠ ا	į		HSFQVENQELSRDKWKVLLNSDFLSHSGA
1295	2645	A	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG
1275	20.5	^^	10133	370	3.0	GRGAPRQEGPGSSWRQV
1296	2646	A	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE
12,0	2040	^	10133		1 331	NIQDVLSALPNPDDYFLLRWLQARSFDLQKS
	İ		]	İ		EDMLRKHMEFRKQQDLANILAWQPPEVVRL
1	{	ļ	Ì	1	ĺ	YNANGICGHDGEGSPVWYHIVGSQDPKGLLL
	}	· ·				SASKQELLRDSFRSCELLLRECELQSQKLGKR
I	1	l	ļ			VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE
1291	2047	^	10138	40	407	TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT
i	<b>!</b>	Į.		•	1	CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC
1	]	}		J	}	
1000	2649	<del>  </del>	10156	04	452	QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG
			1			QEMYLRFDQTTRRSPYRMSRILARHQLVTKI
		1	1			QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP
		<u> </u>	<u> </u>			INIVAVKNDHDFLEKDLGEPLCRRLNT
1299	2649	A	10161	[ 1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSQPT
	Į.	İ	l	[		ASAQLENMEEAPKRVSLALQLPEHGSKDIGN
				1	1	VPGNCSENPCQNGGTCVPGADAHSCDCGPGF
1	1	ł	1	i		KGRRCELACIKVSRPCTRLFSETKAFPVWEGG
		1	<b> </b>	<u> </u>		VCHHV
1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA
	<u> </u>	1	1		ļ	VTYTGILGPGLIGNILALWVFYGYMKETKRA
I	1	1	[		1	VIFMINLALADLLQVLSLPLRIFYYLKHDWPF
	1	Ì	1		1	VPV
1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE
				1 -		KQLQCMPMEGRGRASSSISDLQGKGFEKGTG
1	l	l	1	[	1	EKHVPGVGSARHSPQASAGGSPWQRGKAQT
		1	1	1		RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR
		1		1	1	LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP
	1	1	1			PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA
			· ·			LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP
	}	}	j	J	1	RTLSVEEPGVECNQLCLYADVTDPVLCLGQK
	[	1	l	}	Ī	I
	l	l	<b>]</b>	J		DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP
İ						ARRLSESLHVVDENKNESKIEREHKRRTSTPV IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID NO: of NO: of No: of nucleotide eotide sequence wence where the peptide eotide sequence w	cteine
nucl- eotide seq- uence  in	Sieme,
eotide sequence 09/496 09/496 ocorresponding of to last amino acid residue of peptide residue of peptide sequence 0	ne
sequence    Sequence	110,
uence  914  ng to first amino acid residue of peptide residue of peptide sequence  916  917  1 T-Thronine, V=Valine, W=Typtophan incleotide insertion  8 STLKNEKHLKKDDSETPHLKSILK KEKPEREKTYSEDKLSVKHYKGDD DETELHSSEKGLK VEENIOKQSQOTT TERKSKHRNERKLSVLGKDGKPVSE NVRKENNKKERRLSAETKAEHKSS IQKDSLGSKQHGTLQRRSESYSEDK NMDSNLKPEEVVHKEKRTKSLLEE KSKTQKQVKVVETELQBATKQA KEKNTEENDSEKQRKSRVEKKPFEE LETASSSAAHSTQKDSSHRAKLPLAKL KDSTSTRLERKLSDGHKSRSLKHSSE ENKSDDKDGKEVDSSHEKARGNSSS SRRLCENRRGSLSQEMAKGEKLAA-  9 SOSSLQRPKKSGDMT.UPPQEPPMEID NVFEVSKTQDNRNNNSHQDIDSENA ATVQKDELRTCTADSKATAPAYKPE NNSSKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASTKHIPBAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKKTVENGKS HVVGLNTEKY AETVKLKHKRSPGK VERRIENSEVDTSAGSGAPSVLHQ DVATGPRAEKTSVATSTEGKDKDAV KCDGSLSRDSEIVEGTITIESVESDG GFASETFLTSTKGESGAPASCEGGGST LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITIESVESDG TEIRRAGSISSEEVVEGSGOGNMMRMMP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSS GESAVTSTGTIEDGEGPASCTGSEDS	ne.
amino acid residue of peptide residue of peptide peptide sequence  T=Threonine, V=Valine, W=Tryptophan Y=Tyrosine, X=Unknown, Y=Stop codo (>possible nucleotide deletion, \possible nucleotide deletion, \possible nucleotide insertion  KSTLKNEKHLKKDDSETPHLKSLLK KEKPEREKTPSEDKLSVKHKYKGDK DETELHSSEKGLK VEENIQKQSQOTI TERKSKHRNERKLSVLGKDGKPVSS NVRKENNKKERRLSAAKTKAEHKSI IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEEVVHKEKRRTKSLLEH KSKTQGKQVXVVETELQEGATKQA KEKNTEENDSEKQRKSKVVETELQEGATKQA KEKNTEENDSEKQRKSKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSAAHSTQKSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI SRRLCENRRGSLSQEMAKGEEKLAA SQSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHARAFAKPA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHADIDSENN ATVQKDELRICTADSKATAPAYKPA NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQQDATLEHST LSSVTVVPLRESYDPDVIPLEDKRTY STSPADHSALPVQSLTVRESEVLKTS STSPADHSALPVQSLTVRESEVLKTS STSPADHSALPVQSLTVRESEVLKTS STSPADHSALPVGTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRIENSEVDTSAGSGSRSVLHQ DVATGRRARKTSVATSTEKKDKDDAV KCROSLSTROSEVSLTSGSGECAVAESE LAVHAVKIEANVISVVTEEKDDAV KCDGSLSRDSEVEGTTFISEVEDDAV KCDGSLSRDSEVEGTTFISEVEDDAV KCDGSLSRDSIPVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSS GESAVTSTGITEDGEGPASCTGSEDS GESAVTSTGITEDGEGPASCTGSEDS	10,
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Siop codor /=possible nucleotide deletion, \text{v=possible possible nucleotide deletion, \text{v=possible nucleotide deletion, \text{v=possible possible possible possible possible publication} nucleotide deletion, v=possible possible	,
Peptide   /=possible nucleotide deletion, \r=possible nucleotide insertion   rucleotide insertion   RSTILKNEKHLKKDDSETPHLKSLLK   KEKPEREKTPSEDKLSVKHKYKGDC   DETELHSSEKGLKVEENIQKQSQQTT   TERKSKHRENEKLSVLGKDGKPVSE   NVRKENNKKERRLSAEKTKAEHKSI   IQKDSLGSKQHGITLQRRSESYSEDK   NMDSNLKPEEVVHKEKRRTKSLLE   KSKTQGQVKVVETELQEGATKQA   KEKNTEENDSEKQRKSKVEDKPFEE   LETASSSAHSTQKDSSHRAKLPLAKI   KDSTSTRLERKLSDGHKSRSLKHSSI   ENSDDKDGKEVDSSHEKARGNSSI   SRRLCENRRGSLSQEMAKGEEKLAA   SGSSLQRPKKSGDMTLPEQEPMEID   NVFEVSKTQDNRNNNSHQDIDSENM   ATVQKDELRTCTADSKATAPAYKPC   NSNSEKHADHRSTLTKKMHQSAVS   KEPIHRGTTEVNIDSETVHRMLLSAF   QKNLKNTAAEEHVAQGDATLEHST   LSSVTVVPLRESYDPDVPLFDKRTV   STSPADHSALPNQSLTVRESEVLKTS   EGFTVDTPAKASITSKRHIPEAHQAT   GKVIMPLGSKLTGVIVENERITIKEGG   KKENDLNAEPNLKQTIKATVENGKF   HVVGLNTEKYAETVKLKHKRSPGK   VERRNENSEVDTSAGSGASPSVLHQ DVATGPRRAEKTSVATSTEGKDKD   AGPATTTSSETRQSEVALPCTSIEADI   HSRNNPLHVGAEASECTVFAAAEEG   GFAESETFLTSTKEGESGECAVAESE   LAVHAVKEANNNSVVTEEKDDAV   KCDGSLSRDSIVEGTITTISEVSEDS   TEIRAGSISSEEVDGSQONMMRMGP   TVTCTGAEGRSDNFVCSVTGAGPRI   GAGVVLGDNDAPPGTSASQEGGSS   GESAVTSTGITEDGEGPASCTGSEDS	
Sequence   nucleotide insertion	
KSTLKNEKHLKKDDSETPHLKSLLK KEKPEREKTPSEDKLSVKHKYKGDO DETELHSSEKGLKVEENIQKQSQOT TERKSKHRNERKLSVLGKDGKPVSE NVRKENNKKERLLSAEKTKAEHKSI IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEEVVHKEKRTKSLLEE KSKTQGKQVKVVETELQEGATKQA KEKNTENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAK KDSTSTRLERKLSDGHKSRSLKHSSE ENKSDDKDGKEVDSSHEKAGRNSSI SRRLCENRGSLSQEMAKGEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENA ATVQKDELRTCTADSKATAPAYKP NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENTKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRGSVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVYTEEKDDAV KCDGSLSRDSIVEGTITTISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAGGRSDNFVICSVTGAGPRI GAGVVLGDNAPPGTSASQEGDGS	C
KEKPEREKTPSEDKLSVKHKYKGDO DETELHSSEKGLKVEENIQKQSQQTI TERKSKHRNERKLSVLGKDGKPVSE NVRKENNKKERRLSAEKTKAEHKSI IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEVVHKEKRRTKSLLEI KSKTQGKQVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAAISTQKDSSHEKARGNSSI SRRLCENRGSLSQEMAKGEEKLAA KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENTIKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSIVGCTITTISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAGGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGS GESAVTSTGITEDGEGPASCTGSEDS	VEVICE
DETELHSSEKGLKVEENIQKQSQQTI TERKSKHRNERLSVLGKOGKPVSE NVRKENNKKERRLSAGKTKAEHKSI IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEEVVHKERRITKSLLEE KSKTQGKQVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDDKDCKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKSGDMTLIPEQEPMEID NVFVSKTQDDRNINNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPC NSNSEKHADHRSTLTKKMHIQSAVS KEPHRGTTEVNIDSETVHRMLSAS QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVYLLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDN AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMRMMGP TVTCTGABGRSDNIFVICSVTGAGPRI GGAGVULDNDAPPGTSASOEGDGST	
TERKSKHNERKLSVLGKDGKPVSE NVRKENNKKERRLSAEKTKAEHKS) IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEEVVHKEKRRTKSLLEE KSKTQGKQVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAK KDSTSTRLERKLSDGHKSRSLKHSSE ENKSDDKDGKEVDSSHEKARGNSS) SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHQSAVS KEPHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLITGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEAD HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITPISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GGGVVLGDNDAPPGTSASOEGDGS	
NVRKENNKKERRLSAEKTKAEHKSI IQKDSLGSKQHGITLQRRSESYSELDE NMDSNLKPEEVVHKEKRRTKSLLEE KSKTQGKQVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEOEPMEID NVFEVSKTQDNRNNNSHQDIDSENA ATVQKDELRTCTADSKATAPAYKP NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAP QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTY STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEAD HSRNNPLHVGAEASECTVFAAAEEC GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQONMMRMOP TVTCTGAEGRSDNFVICSVTGAGPRI GGGVVLGDNDAPPGTSASQEGDGSI GESAVTSTGITEDGEGPASCTGSEDS	YESSDOK
IQKDSLGSKQHGITLQRRSESYSEDK NMDSNLKPEEVVHKEKRRTKSLLEE KSKTQGKQVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSEM ATVQKDELRTCTADSKATAPAYKP NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVYTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMRMMGP TVTCTGAEGRSDNFVICSVTGAGPR GGGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
NMDSNLKPEEVVHKEKRRIKSLLER KSKTQGKVKVVETELQEGATKQA KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAK KDSTSTRLERKLSDGHKSRSLKHSSR ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENA ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAA QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKR HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKEANVNSVYTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGS GESAVTSTGITEDGEGPASCTGSEDS	
KSKTQGKQVKVVETELQEGATKQA KEKNTEEDDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKF HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGS GESAVTSTGITEDGEGPASCTGSEDS	
KEKNTEENDSEKQRKSKVEDKPFEE LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSE ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNIKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAFSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEAD HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGPP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
LETASSSAHSTQKDSSHRAKLPLAKI KDSTSTRLERKLSDGHKSRSLKHSSI ENKSDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPP NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNIKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENOKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTTTFISEVESDG TEIRAGSISSEEVDGSQONMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	TCYTCH
KDSTSTRLERKLSDGHKSRSLKHSSE ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENM ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGE KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDK AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITTISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	IGVEPV
ENKSDDKDGKEVDSSHEKARGNSSI SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKF HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
SRRLCENRRGSLSQEMAKGEEKLAA SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITIFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
SGSSLQRPKKSGDMTLIPEQEPMEID NVFEVSKTQDNRNNNSHQDIDSENN ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPR GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
NVFEVSKTQDNRNNNSHQDIDSENM ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST. LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGO KKENDLNAEPNLKQTIKATVENGKE HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESS LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
ATVQKDELRTCTADSKATAPAYKPO NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST. LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGO KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEAD! HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
NSNSEKHADHRSTLTKKMHIQSAVS KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
KEPIHRGTTEVNIDSETVHRMLLSAF QKNLKNTAAEEHVAQGDATLEHST LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
QKNLKNTAAEEHVAQGDATLEHST. LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKR HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
LSSVTVVPLRESYDPDVIPLFDKRTV STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
STSPADHSALPNQSLTVRESEVLKTS EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
EGFTVDTPAKASITSKRHIPEAHQAT GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV: KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
GKVIMPLGSKLTGVIVENENITKEGG KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV: KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
KKENDLNAEPNLKQTIKATVENGKK HVVGLNTEKYAETVKLKHKRSPGK VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	LLDGKQ
HVVGLNTEKYAETVKLKHKRSPGK: VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATITSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV: KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	iLVDMA
VERRNENSEVDTSAGSGSAPSVLHQ DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
DVATGPRRAEKTSVATSTEGKDKDV AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
AGPATTTSSETRQSEVALPCTSIEADI HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
HSRNNPLHVGAEASECTVFAAAEEG GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	VILSPVK
GFAESETFLTSTKEGESGECAVAESE LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
LAVHAVKIEANVNSVVTEEKDDAV KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
KCDGSLSRDSEIVEGTITFISEVESDG TEIRAGSISSEEVDGSQGNMRRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
TEIRAGSISSEEVDGSQGNMMRMGP TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
TVTCTGAEGRSDNFVICSVTGAGPRI GAGVVLGDNDAPPGTSASQEGDGSV GESAVTSTGITEDGEGPASCTGSEDS	
GAGVVLGDNDAPPGTSASQEGDGST GESAVTSTGITEDGEGPASCTGSEDS	
GESAVTSTGITEDGEGPASCTGSEDS	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
SESEENGESAMDSTVAKEGTNVPLV	
DEGIVTSTGAKEEDEEGEDVVTSTGI	
ASTCTGLGEESEGVLICESAEGDSQI	
VEAEAGAAIMNANENNVDSMSGTE	
DICSSAKGIVESSVTSAVSGKDEVIP	
GPMTSAASDQSDSQLEKVEDTTISTC	
YDVLVSGEVPECEVAHTSPSEKEDE	DIITSVE
NEECDGLMATTASGDITNQNSLAGG	
VLIISTSTTNDYTPQVSAITDVEGGLS	DALRTE
ENMEGTRVTTEEFEAPMPSAVSGDD	
RSEEKDECAMISTSIGEEFELPISSAT	
LQPVAAAVEERATGPVLISTADFEGH	
EAESPLASTSKEEKDECALISTSIAEE	
GVVVESENERAGTVMEEKDGSGIIST	
EGPVSSAVPQEEGDPSVTPAEEMGD	
TSEGCEAVMIGAVLQDEDRLTITRVI	EDLSDA
ALISTSTAECMPISASIDRHEENQLTA	DNPEGN
GDLSATEVSKHKVPMPSLIAENNCR	CPGPVR
GGKEPGPVLAVSTEEGHNGPSVHKP	SAGQGH
PSAVCAEKEEKHGKECPEIGPFAGRO	GOKESTL
HLINAEEKNVLLNSLQKEDKSPETGT	CAGGSST
ASYSAGRGLEGNANSPAHLRGPEQT	SGQTAK
DSSVSSIRYLAAVNTGAIKADDMPPV	
EHSFLPAEQQGSEDNLKTSTTKCITG	/QGTVA

C050 10	OFO ID	34-4	000	Dandistad	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Mct	SEQ	Predicted	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nuci-	peptide	j	in			I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	'	· {	914	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
<u> </u>			ļ			Y=Tyrosine, X=Unknown, *=Stop codon,
1				residue of	sequence	/=possible nucleotide deletion, \=possible
<b>!</b>				peptide		
				sequence		nucleotide insertion
<u> </u>						SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK
						WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS
<b>(</b>		l i				SEENVCDIGNEESPLNVLGGLKLKANLKMEA
1						YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE
!						PLLVNESLNVENSGFRTNEEIHSEŞYNKGEISS
		1	i			GRKDNAEAISGHSVEADPKEVEEEERHMPKR
]						KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC
1			· 			PETEPHATKEENSRDLEELPKTSSETNSTTSRV
					İ	MEEKDEYSSSETTGEKPEQNDDDTIKSQE
1302	2652	A	10167	321	842	EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH
						FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG
1						NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE
1		}				DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP
1						EARETKCYVRSSVGCVEPLTTQAEVTENLDR
				ļ		KNSQQVFKLLKKK
1303	2653	A	10171	206	429	NMILLKKRRLLINSLGEGTINGLLDELLETNV
1505	2033	1 **	101/1	200	"2"	LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA
				i		RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF
1304	2034	^	10104	370	1324	YHEAVVLFTQALKLNPQDHRLFGNRSFCHER
ľ						LGQPAWALADAQVALTLRPGWPRGLFRLGK
ļ					Ì	ALMGLQRFREAAAVFQETLRGGSQPDAAREL
ĺ		[		1		RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA
				•		ELAPSGLPSLRCPRSTALRSPGLSPLLH
		ļ.,i	10101	<u> </u>	1 204	
1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ
1	i	j		}		SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR
J				ļ		DFISVTLSFGESYDNSKSWRRRSCWRKWKQL
					İ	SRLQRNMILFLLAFLLFCGLLFYINLADHWKG
					<u> </u>	IRNTCT
1306	2656	A	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY
1						NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI
}	1	ł			l l	GIDDEDDSTFTITVDQKTFHFQARDADEREK
1				<u> </u> -		WIHALEETILRHTLQLQVRVFTWFPDSSLVGA
j				]		FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
		l				DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS
ł	ł	1	} 	ł		SPDQQNYTKSR
1308	2658	A	10214	2	453	ECGGIRQPGPPPPALASAPAATMNRVGGSPS
]	]	1		1	ļ	AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP
	ļ	1		1		SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL
	]					DNITQVMSLHTQYLESFLRSQFYMLRMDGPL
	· ·	Ī		[		PLPYRHYIAIMAAARHQCSYLINM
1309	2659	A	10233	45	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
1303	2039	^	10233	75	121	RALESRTFQGSERSRWGPPLESTKENVQCGH
1	<b>!</b>	1	ł	1	1	RPAFPNSSWLPFHERLQVQNGECPWQVSIQM
	1		ł	1		SRKHLCGGSILHWWWVLTAAHCFRRTLLDM
J .	J	1	1	1		AV
1212	0660	<del> </del>	10041	242	142	AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK
1310	2660	A	10241	243	442	
1	1		1	[	[	HKKGQSAEIQKKRTRRAFKFQRAITGASLADI
L	L	<u> </u>		L	<b></b>	MAK
1311	2661	A	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ
	1	1		1		VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ
	l	ļ	1			LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
	[	1	[	1		QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI
-		ļ ·		-]-	<b>+</b>	DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML
1	l	ł	1	1	1	QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI
1		<b>!</b>		}	1	KAG
1312	2662	Α	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP
1	1	]	1 .	1		SMTILDKKDGEQAKALFEKVRKFRAHVEDSD
			I			<u> </u>

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF
						EHVCKPKVEHLIGYEVFECTHNMAYMLKKL LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	A	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRGRSRSYSRSRSRSWSKERLRERDRD RSRTRSRSTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPV NLRPPVPPPGPLPPSLPPVTGPPPPLPPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	Α	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

ſ	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
ſ	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
-1	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
- }	eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
1	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
.	uence	40.100		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
- (	dence		ì	717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ı			ł				Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	ļ	residue of	sequence	
- 1					peptide	1	/=possible nucleotide deletion, \=possible
-					sequence		nucleotide insertion
ľ						}	NEKLVDEKTILETSFHQHRERAEQLSQENEKL
-1			ł	l	ł	ļ	MNLLQERVKNEEPTTQEGKHELEQKCTGILE
1			1	l	ļ		OGRFEREKLLNIQQQLTCSLRKVEEENQGAL
1				j			EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1			{	(	ſ	Í	EECRVTLEGLKMENGSLKSHLQG
ŀ	1010	2669	<del> </del>	10202	222	070	GECFIMAAVVQQNDLVFEFASNVMEDERQL
J	1318	2668	] A	10303	333	879	GECHNIAA V VQQINDL V FEFASIN V MEDERQL
1			ì	i	ŀ		GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
١				1	1		NDMITESSLDVAEEEIIDDDDDDDITLTVEASCH
- 1		1	ł	1	ĺ		DGDETIETIEAAEALLNMDSPGPMLDEKRINN
-1				1	Ì		NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
-		ŀ					QVQEKYADSPGASSPEQPKRKKK
-	1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLLIIATALSL
-	1317	2007	^	10322	107	""	LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR
- 1			1		[	1	LIUMNOLIMOV VLEMIOU I COMPONIVEDAN
		ļ	I	}	1	}	LPRTLAGLLAGGALGLAGALMQTLTRNPLAD
			1		1	1	PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
- 1			1	}	1	1	MAFAGALVASLIVAFTGSQGGGQLSPVRLTL
- 1		}	ł	ł		1	AGVXL
ŀ	1320	2670	A	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
i	1520	2070	1 ''	10323		1 -	AVVDIQSDKAANVAQEINAEYGESMAYGFG
ı		Ì	(	ĺ		Ī	ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
ı		ļ	į.	ļ			
J		}	1	j	į.		AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
- 1			<u> </u>	<u> </u>	L	L	FSRLMIRDGIQGRIIQINSKSDE
- {	1321	2671	A	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
ļ		i	ļ	1	İ	1	AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
ł		i	1	l	I	l	ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
		ł	İ	l	1		YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
Ì			1	l.		1	MALLFEHRYSVDLLPFVQKAPTDSEA
- 1	1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
- 1	1322	2072	^	10333	23	423	ERKMRAHQVLTFLLLFVITSGASENASTSRGC
				1	1		
				1			GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
		1	1		•		LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL
1	· _		1				LGTLG
- [	1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
- 1			ł				QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
J			1	J	]		YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI
-		1		1		1	LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
-	1324	2674	A	10336	<u> </u>	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
-	1324	2074	1	10330	1 *	1 752	NSVTHHEVKCQGKPLAGIYRKREEKRNAGN
1			1'	1			AVRSAMKSEEOKIKDARKGPLVPFPNQKSEA
			1	1			
4			1	i	1		AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
		1	1			1	PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
ı		ļ	1	1	]		LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
			1	1	ļ	ļ. ·	VIATKQFSRGDFVVEYHGDLIEITDAKKREAL
			[	1	1		YAQDPSTGCYMYYFQYLSKTYCVDATRETN
ļ		1	1	1		1	RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
		ł	1		i	1	RDIAAGEELLYDYGDRSKASIEAHPWLKH
	105-		<del> </del>	1,000	<del> </del>	000	
	1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
			1				RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
			1			1	PMRKSYRGDREAFEETHLTSLDPVKQFAAWF
	1	}	1	1		1	EEAVQCPDIGEANAMCLATCTRDGKPSARML
		1	1				LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
			1		1		VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
				1	1	1	RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
i		1	1	1		1	QLYQDQEVPKPKSWGGYVLYPQVMEFWQG
		ŀ	1				
		1	1	1	1	+	QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
		<u> </u>	<u></u>		<u> </u>	<u> </u>	DWLYERLAP
į	1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
			1	1		1	LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT
			1		1		LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA
				1	L	<del></del>	

SEQ ID	SEQ ID	Met	SEQ	Predicted .	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		Ī	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}	!	}	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}	1	}	1	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		<del></del>	<del> </del> -	sequence	<del></del>	nucleotide insertion HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD
1		)		)		
			İ		•	EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY
}	j	,	}			LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA
		l			'	ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ
		ĺ		ļ		ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD
j	}		[	j		MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
			}			LDKLGYSQKMRFTY
1327	2677	A	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
102.	-0	1.	103.5	( •	300	RYHRLSDMSMLAERRRKQKWAVDPQNTAW
1						SNDDSKFGQRMLEKMGWSKGKGLGAQEQG
1				i		ATDHIKVQVKNNHLGLGATINNEDNWIAHQ
1	ļ			'		DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
ł	i		1			LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL
			·			DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF
}		ļ				TIQEYFAKRMAALKNKPQVPVPGSDISETQVE
						RKRGKKRNKEATGKDVESYLQPKAKRHTEG
		•		_		KPERAEAQERVAKKKSAPAEEQLRGPCWDO
						SSKASAQDAGDHVQPA
1328	2678	Α	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI
1	:					CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
			L			HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	Α	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL
						SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN
1			1			LSFADICVTSTTIPKMLMNIQTQNKVITYIACL
1						MQMYFFILFAGFENFLLSVMAYDRFVAICHP
1		ļ				LHYMVIMNPHLCGLLVLASWTMSALYSLLQI
						LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF
			ļ			LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH
i l						AISSAQGKYKAFSTCASHLSVVSLFYGAILGV
1						YLSSAATRNSHSSATASVMYTVVTPMLNPFI
1220	2600	4	10000	24	0.550	YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	Α	10352	34	2573	IPFLKSCCCCLFDFPPPPLDQVQEEECEVERV
1						TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE
1						TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL
1						FYTERAHVRTLKVLDQVFYQRVSREGILSPSE
						LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS
1						VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ
1						PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR
1						LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT
						EREKVKKAADHCRQILNYVNQAVKEAENKQ
						RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV
						LLQKQDDRLVLRCHSKILASTADSKHTFSPVI
						KLSTVLVRQVATDNKALFVISMSDNGAOIYE
				İ	1	LVAQTVSEKTVWQDLICRMAASVKEQSTKPI
						PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL
					' '	QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD
1				ľ	·	LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
						HLPVSEERWALDALRNLGLLKQLLVQQLGLT
[				¦	l	EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE
						NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE
					ļ	SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
						MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
					j	GDGAVNKEEKDVNLRISGNYLILDGYDPVQE
						SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP
] [						QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI
					ĺ	QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE
						SYTILCORLAGSALTDKHSDKS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end a nucleotide location corresponding to last amino acid residue of peptide sequence	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1331	2681	A	10353	1	2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEG AAGQQPTAPDKSKETNKTDNTEAPVTKIELLP SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP IIMGANIGTSITNTIVALMQVGDRSEFRRAFA GATVHDFFNWLSVLVLLPVEVATHYLEIITQL IVESFHFKNGEDAPDLLKVITKPFTKLIVQLDK KVISQIAMNDEKAKNKSLVKIWCKTFTNKTQ INVTVPSTANCTSPSLCWTDGIQNWTMKNVT YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL ASPGNALRSSLQIALCHFFFNISGILLWPPIPFT RLPIRMAKGLGNISAKYRWFAVFYLIFFFLIP LTVFGLSLAGWRVLVGVGVPVVFIIILVLCLR LLQSRCPRVLPKKLQNWNFLPLWMRSLKPW DAVVSKFTGCFQMRCCCCCRVCCRACCLLC GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET FDNITISREAQGEVPASDSKTECTAL
1332	2682	A	10354	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPR GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP TSDLGEIHNWTELLDLFNHTLSECHVELSQST KRVVLFALYLAMFVVGLVENLLVICVNWRG SGRAGLMNLYILNMAIADLGIVLSLPVWMLE VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF LVCLSVDRYVTLTSASPSWQRYQHRVRRAM CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM APFETYSTWALAVALSTTILGFLLPFPLITVFN VLTACRLRQPGQPKSRRHCLLLCAYVAVFV MCWLPYHVTLLLLTLHGTHISLHCHLVHLLY FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP TQPLTPS
1333	2683	A	10358	2	884	AAGAGADGREPASERASRAEPPAVAMGQND LMGTAEDFADQFLRVTKQYLPHVARLCLIST FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF GIIALQTIAYSILWDLKFLMRNLALGGGLLLL LAESRSEGKSMFAGVPTMRESSPKQYMQLGG RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI LVAIGFKTKLAALTLVVWLFAINVYFNAFWT IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL GPGGVSMDEKKKEW
1334	2684	A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFP ELPLPHVPGQESAKRRSARRFLIMSELTKELM ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS WRDCSQEEQKELLCHTLCDILESACCDHSGS YCLVSWLRGKTTEETASISGSPAESSCQVEHS SALAVEELGFERFHALIQKRSFRSLPELKDAV LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN EIEDASEPLIDPVYGHGSQSLINLLLTGHAVSN VWDGDRECSGMKLLGIHEQAAVGFLTLMEA LRYCKVGSYLKISKIPYLDCLASETHLTVFFA KDMALVAPEAPSEQARRVFQTYDPEDNGFIP DSLLEDVMKALDLVSDPEYINLMKNKLDPEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL
1335	2685	A	10375	82	2929	KQSNYNEKVMYVEGTAVVMGFEDPMLQTD DTPIKRCLQTKWPYIELLWTTDRSPSLN  TRTKRRLGREKAMASPPRGWGCGELLLPFML LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG
1336	2686	Ā	10379	1	557	GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNAAGKRDGKAPAGGNGN KKKSGKKEKK  RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK LAKLQAQVRIGGKGTARRKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK
1337	2687	A	10380		1263	PEDIDEEDDDVPDLVENFDEASKNEAN  IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QPN
1339	2689	Α	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690		10388	-	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRICSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTIVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

NO. of nucleotide coide sequence unnecedible sequence of the control of the contr	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence unoce unoce unoce corresponding to general process of the pro							D=Aspartic Acid, E=Glutamic Acid
Double   D							
pequence    1914						corresponding	
uence   914   ng to first amino acid residue of peptide peptide peptide peptide sequence   7—Thronine, V-Valine, W-Prytophan, Y-Prytoph	1		ļ	l .	correspondi		, , , = = : :,
mimo acid residue of peptide sequence peptide sequence of peptide sequence peptide sequence peptide sequence peptide sequence per company to the sequence per company to t			1				
residue of peptide sequence     Y=Tyrosine, X=Unknown, **Sign codon, *possible nucleotide diestrion, *possible nucleotide diestrion, *possible nucleotide diestrion, *possible nucleotide diestrion. *possible nucleotide nucle			[			Ĭ	
peptide			1				
Inuclotide insertion	(		ĺ				
SQVGGKRFECKDCGETTNISAALAEHRKIHA RGYLVECKNQECEAPMPSITTSELGKIYGK DKFYECKVCKETFLHSSALEHQKHRGDDKD NEREHERERERREGTERSPEAMPSETGKMYGK DKFYECKVCGETFLHSSALEHQKHRGDDKD NEREHERERERREGTERSPEAMPSETGKMYGK KEMMPCKVCGETFLHSSSLKEHQKHRGDDKD PENKIG VCEETTFUGGSLKRRGKYNKEKLC DFTDGRDAFNQSSLESHQKHRSKNLFFGR GYEKSVHSQFTESGKSHTITRPLESDEDEKA FTISNPYENOKIPTKENVVEAKSVERSVHISL ASVEAQKSHVAGPTESGKSHTITRPLESDEDEKA FTISNPYENOKIPTKENVVEAKSVERSVHISL ASVEAQKSHVAGPTESGSFDAN HQRVEAGGNTISGGREYSRSVHISL VASKPPDS HYONGLEVENBEGGESSKYTSDLDNRGKDFAR ENPCEGGSKNRHYEDSVIGSVPRAKEDGKSPA ENPCEGGSKNRHYEDSVIGSVPRAKEDGKSPA GEGGGFKKDGFEFSPSPSNVERVGKARAKK YUEHRSNETSVHISLPPGEOTFRRGMLYBCO ECGCF-AFSDLTEHQKHDEEKPSGSENYE WSVIRSLAFTDQTSVAGEOVAKBQARNCK VEHRSNETSVHISLPPGEOTFRRGMLYBCO ECGCF-AFSDLTEHQKHDEEKPSGSENYE WSVIRSLAFTDQTSVAGEOVAKBQARNCK DPROFFATSEDLNTNGKYTDQESSHGEBGGE NTDGEETHSEETHGGETEDPVIQGSDMEDPQ KDPDDKTYECEDGOLGFVDLTDTI-DHQKYH SRKCLVDSRRYTHSVHIFTHSISESYORDYTCEG LYCPKCGSEFHISSFLEFHQRIEEDOLJ SNM KGCDDGFLALLPMKPRRNRAAERVPALAGES QMAERAHFGLALTERQRSOTFERH PLCAVCG SEFNYAPAELADHVYTHVKNDEYPYSTGSYNTITS FLTEPLKGARFYECKLOCKSFHISTVLTKHEE LHLEEGEEDBAAAAAAAAAAACAPURAHAYHVHV VVLRIQGLINVEAAEFEVEAAEFEVEAAEFEVE GAABPPECKAEFEVEAAEFEVEAAEFE GAADADPDAVGEDPEGEEDOLGIOVEEPTYDC HECTETTTSSTAFSEHLKTHASSMIERPANAFG ECGYTERTAMPSTELEFANAFG ECGYTERSTAFSEHLKTHASSMIERPANAFG ECGYTERTAMPSTELEFANAFG ECGYTERTAMPSTELIGGROUPERSTELL QUANCDMEDERSPENTAMPSTELLIGGNOSTICK VCNAMPSTELLOFTONOVAGAAAAAAACHAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1						nucleotide insertion
RGYLVECKNOECEEAPMPSTYTESLOKIYOK DKYYECKVCKETFILSSALEHOKHRIDOND NEREHERERERERGETERFSPALNEFOKMYG KERMYECK VCGETFILSSALEHOKHRIDOND NEREHERERERGETERFSPALNEFOKMYG KERMYECK VCGETFILSSALEHOKHRITGN PFENKGK VCEETFIPGQSLKRROKTYNKERLC DFTDGRAFMQSSELSHOKHRISKNIHFOR GYEKSVHISGPTESOKSHITTRPLESDEDEKA FISSNYPENOKLIPTKENVYZAKSYERSVHISL ASVEAQKSHSVAGPSKOKSHITTRPLESDEDEKA HORVER AGONTSGEREKSVSHISL VASKPPRS HNONEL VESNEKGESSIYISDLNDKROKUPAR ENPCEGGSKRRYNYERSVSKSVHISL VASKPPRS HNONEL VESNEKGESSIYISDLNDKROKUPAR ENPCEGGSKRRYNYERSVOKSYPRAKYOKSVP GEOSGEFKKUDGEFSVPSSNYREYOKARAKK YIEHRSNETSVHISLPFGEOTFRRRGMLYECQ ECOLET AHSSDLTEHOKHIDOLEPOSGSNYT WSVIRSLAYTDPQTSYAQEQYAKEQARINGK DFROPATSBEDLNINGKYTYDQESSHORD NTDGETHISETHOQHIDDREPOSGSNYT WSVIRSLAYTDPQTSYAQEQYAKEQARINGK DFROPATSBEDLNINGKYTYDQESHORD NDFORTSBETHOQHIDDREPOSGSNYT WSVIRSLAYTDPQTSYAQEQYAKEQARINGK DFROPATSBEDLNINGKYTYDQESHORD NDFORTSBETHOQHIDDREPOSGSNYT WSVIRSLAYTDPQTSYAQEQYAKEQARINGK COLOGGARASELIPGLALTEPOSGSTORT WSVIRSLAYTDPQTSYAQEQYAKEQARINGK SKCLVDSREYTHSVHITHSISETYORDYTICEQ LYCCYKCGSEHISSTCHEFOKHEQDLYTOC LYCCYKCGSEHISSTCHEFOKHEQDLYTOC GENERALDPOTATION STANDAY KGCDDGFIALLPMRYRNRAARENPALAGSA RKCLCOGGFHISSALHHAHRAFINLINGKY GCDDGFIALLPMRYRNRAARENPALAGSA RKCLCOGGFHISSALHHAHRAFINLINGKY KGCDDGFIALLPMRYRNRAARENPALAGSA RKCLCOGGFHISSALHHAHRAFINLINGKY GCDGGFHISSALHHAHRAFINLINGKY KGCDDGFIALLPMRYRNRAARENPALAGSA RKCLCOGGFHISSALHHAHRAFINLINGKY GCDGGFHISSALHHAHRAFINLINGKY GCDGGFHISSALHAHRAFINLINGKY VIRIGGLINVEAAEPEVEAAEPEVEAAPEVE EAAFPROEAGEDGOEGAAFOCAGORONGAE QPNGGADEPDGAGEDDFERRAFEECKAEPE GADADPOVOGEDFEGEOGOGIOVEPTYDC HECTETTISTSTASBEHLKTHASMIFFEANAFG ECGYUERASTSTGGAAAAAGQEVARTIGNSA ASRRSPAARPPVAPPALPRGRGTGGSTSLS APAVLVVAVAVVVVVVVVVVAVAVAMANYHV PRGSFEVYKLAVVAVAVVVVVVVVAVAVAMANYHV PRGSFEVYKLAVVAVAVVVVVVVAVAVAMANYHV PRGSFEVYKLAVVAVAVVVVVVAVAVAMANYHV PRGSFEVYKLAVVAVAVVVVVVAVAVAVAVAVVVVXAVAVAMANYHV FKGGOGGAADGEADPKH VCNPAFRLIARQLAKHRAHHAHNGWPKSNI, UKMGCMEBGRIGGTGGTGCTGGALDPKH VCNPAFRLIARQLAKHRAHHAHNGWPKSNI, UKMGCMEBGRIGGTGGTGCTGGALDPKH VCNPAFRLIARQLAKHRAHHAHNGWPKSNI, UKMGCMEBGALDPKH VCNPAFRLIARG					0040000		
DIKFYECKYCKETFLHSSALIEHQKHIFGDIXD NREHEREREREREFREPRSPALNEFQKMYG KEKMYECKYCGETFLHSSSLISHQKHITT&GN KEKMYECKYCGETFLHSSSLISHQKHITKGN FERNKGKYCEETFIRGPSALNEFQKMYG KEKMYECKYCGETFLHSSSLISHQKHITKGN FERNKGKYCEETFIRGPSALNEFQSLERBQKTYNKEKLC DFTDGRDAFNQSSELISHQKHISKKNLFFGR GYEKSYHSGYPTENGSLKRRICKYNKEKLC DFTDGRDAFNQSSELISHQKHISKKNLFFGR GYEKSYHSGYPTENGSKHITTRELSEDDEKA FITISNPYENOKIPTKENVYRAKSYERSYHBLI JASVEAQKSHNYGARPKYMARSTIGSDAN HQRVRAGGINTSGGREYSRSVHISL VASKPPPS HNGNEL VESNEKGESSIYISDLNDKRQKJPAR EPPCEGGSRNRNYEDSVJQSYFAKRYGKSYP GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVSSNRVEY GEGSGEFKKDGFSVFSSNRVEY GEGSGEFKKDGFSVSSNRVEY GEGSGEFTSSTLFEHQKHDDEKFSGSKNYEY GEGSGEFKKDFFGFTGFTGFTGFTGFTGFTGFTGFTGFTGFTGFTGFTGF	1		Ì				RGVI VECKNOECEE A EMPSPTESEI OKIVOK
NEREHERERERGETFRPSPALNERGKMYR KERMYEKVCGETTLHRSSLKEHQKHTRIGN FERNGKWYEKVCGETTJIPGOSLKERQKTYNKELL DFDTGDAPAMOSSELSEHQKHTRIGN FERNGKWYEKVCENTDIPGOSLKERQKTYNKELL DFDTGDAPAMOSSELSEHQKHTRENSPERAKER FERNGKAFT AN FIRSWYENQKIPKENYVEAKSYERSVIHSIA ASVEAQKSHSVAGPENEN WARSTIGSTDAND HQRVRAGGNTSEGEYSSVIHALASSIPRS HNGMELVESNEKGESSYINDLNDKRQKTAR ENPECGGKNRNYKEDSVIQSVRAKKKK YEHRSNETSYHBLIPGGTFRPROMLYEC ECCEPTAISBLITEHQKHIDERKYGNSY GEGSGFKKDGEFSVPSNNTER YGKARAKKK YEHRSNETSYHBLIPGGTFRPROMLYEC ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEGSCORD KARCARKKK YEHRSNETSYHBLIPGGTFRPROMLYEC ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEGSCORD KARCARKKK YEHRSNETSYHBLIPGGTFRPROMLYEC ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKHIDERKFGSRNYC WSVIRSLAYTDRQTSYAGEAGNEK ECCEPTAISBLITEHQKATCH WSVIRSLAYTDRQTSYAGEAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							
KEEMYECK VCGETTI-INSSILKEHQKIHTTIGN PFENRGKVCEETPIGGSLERROK/INSERL) PFENRGKVCEETPIGGSLERROK/INSERL) PFENRGKVCEETPIGGSLERROK/INSERL) PFENRGKVCEETPIGGSLERROK/INSERL) PFENRGKVCEETPIGGSLERROK/INSERL) ASVEAQKSHSV-GRESPQK-WHAENSEPER-SUHBLI. ASVEAQKSHSV-GRESPQK-WHAENSEPER-SUHBLI. ASVEAQKSHSV-GRESPK-WAENSEVER-SUHBLI. ASVEAQKSHSV-GRESPK-WAENSEVER-SUBJICATION-IN- HQRVRAGGNTSEGRE-YSRSVIHSL-VASKPPRS HNONEL-VESNEK-GESSIYISDL-INDKROKUPAS ENPCEGGSKRRN-YEDS-VIQSAV-RAKPQKS-YP GEGGGEF-KROGEFS-YFSSNVIRE-YQKARAKK- YEHRESNETS-YHB-SIP-FGGT-TRFRGML-YECQ ECGECF-AHSSDL-TEHQKHERREK-PSGSNNY-E WSVINSLAPTIPOTS-YAQEQ-YAKEQARNICK DFRQFF-ATSEDL-INTOKINTYDQEKSHGEBSG-GBE NTDGEETHSEETHGGSTEEP-YGK-GRESPH-STR- SIRCLL-ON-SPETIAL-TREATING-INTOKINTYDGE-SHGEBS-GBE NTDGEETHSEETHGG-TIE-PYGGBS-MEP-Q KDPPDDKIYT-CED-CGL-GF-VDL-ITD-LTD-HQK- SRC-LUSSRY-THS-WHITHSISSE-YGR-PYTGE-Q LYSC-CROSSFH-SST-LFEHQR-HE-QD-L-YS- SRC-LUSSRY-THS-WHITHSISSE-YGR-PYTGE-Q LYSC-CROSSFH-SST-LFEHQR-HE-QD-L-YS- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YGR-PYTG-L-Y- SRC-LUSSRY-TH-SW-HTH-SISSE-YR- K-GCDG-FILL-SW-HT-SW-HTH-SISSE-YR- K-GCDG-FILL-YB-C-C-G-G-FILL-YB-C-C-G-G-SF-YP-S-SAN-H-R-BR-HA-H-R-BD-LL-SR- R-C-LUSSRY-TH-SW-HT-SH-SW-H-P-S- LYC-LUSSRY-TH-SW-HTH-SISSE-YR- LYC-LUSSRY-TH-SW-HTH-SISSE-YR- LYC-LUSSRY-TH-SW-HTH-SISSE-YR- LYC-LUSSRY-TH-SW-HTH-SISSE-YR- LYC-LUSSRY-TH-SW-HTH-SISSE-YR- LYC-LUSSRY-TH-SW-HTH-SW-HT- SRC-LUSSRY-TH-SW-HT-SW-HT- SRC-LUSSRY-TH-SW-HT-SW-HT- SRC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- SRC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW-HT- LYC-LUSSRY-TH-SW	1		ł				
PFENKGKVCEETFIPGGSLKRROKTYNKELD DFTDGRDAPMGSSELSEHOKHISRKMLFEGR GYEKSVHISOPFTEGOKSHTITRPLESDEDEKA AFTISSNYFENOKITKENVYPAKSYERSVHISUL ASVEAOKSHSVAGPSEPKVMAESTICSTDATM HQRVRAGGNTSEGREYSRSVHISULVASKPPRS HNONELVESNEKGESSIYISDINKRCKIPAR ENPCEGGSKRRINVEDSVIQSVFRAKPQKSY GEGSGFEKKDGESTYSSNNREVYGKARAKKK YIEHRSNETSVHISLFFGEQTFRPRGMLYECQ ECGECFAHSDILTERIQKHIPREKPGSRNYE WSVIRSLAPTDPQTTSYAQEQYAKEQARNKCK DFRQFFATSEDINTNQKIYOQEKGESQGE NTDGEETHSEPTHGOETHEPVIQGSDMEDPQ KDDPDDKIYTECEOCGG.FVVLIDTHOKYM SRKCLVDSREYTHSVHITHSISEVQRDYTGGE LYCCYKCGESHISSSLFHERQHIEDQDLYSM KGCDDGFIALIPMKPRRNRAERRPALAGSA RCLLCGQGHISSSLIFENRHEDQDLYSM KGCDDGFIALIPMKPRRNRAERRPALAGSA RCLLCGGGHISSSLIFENRHEDDLIDGS QMAEEAIPGIALTFPQRSQTEERLFECAVCG ESPVPRAELADHVTVHKNEPYEYGGSYTHIS FILTEPLKGAIPFYECKDCOKSHISTVLTKHKE LHLEESEEDBAAAAAAAAQEVEANVHYPQ VVLRIQGLNVEAAFEVEAAEPEVAAEPEV EAAEPNGEAEGPDGEAAEPGGAGQNOGBA QPNGGADEPDGGAGEDPERGEGDQBIQVEEPYYDC HECTETTFSTSTAFSHLKHTASMIEPRANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHISLARHQNHTHG GCSGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHISLARHQNHTHG GCSGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHISLARHQNHTHG FRAGVILVYAVAVVVVVVSAVAWAMANNITHP PROSEPVEKINTYOQDGEBHERGGALSHLU PROVILVRIYGKKTELLVDRDEEVKSFKUL QAHGCAPQLYCTRVOQCEBALDFIX VCPABIFILLARGLAKHAHAHHANGWPKSNI. WLKMGKYFSLIPTOOTA DEDINKRHYSDDYSSQI LQEEMTWMKELISNLGSVVLCHHDLLCKNII VNEKGADVGTVARYDGAKTSLIGVSSVAG VCPABIFILARGLAKHAHAHANGWPKSNI. WLKMGKYFSLIPTOA DEDINKRHYSDDYSSQI LQEEMTWMKELISNLGSVVLCHHDLLCKNII VNEKGADVGTVARYDGAKTSTERGQHLGGALDPKM VCPABIFILARGLAKHAHAHANGWPKSNI. WLKMGKYFSLIPTOA DEDINKRHYSDDYSSQI LQEEMTWMKELISNLGSVVLCHHDLLCKNII VNEKGADVGTVARYDGAVTTITAYYRGAMGIC LANDINGAYSTEPDFLGVAVRRNQYFKM KPEVYLLKVPR EKGGGTEVTEEKVELIFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTSVTEEKVELIFFQNONGALASHFYW FRAGGTTSTEEKVELIFFQNONGACHAGALGGE LMTDTINEESPAAVQDWSTGKTTYSWDNAQ VILVONKCOMEDERVISTERGQHLGGE		,	ŀ				l
DFTDGRDAPMQSSELSEHQKHSRKNLTEGE GYEKSVHISGPTEQQKSHTTRPLSDEDEKA FTISSNPYENQKIPTKENVYEAKSYEESVHBLI ASVEAQKSHSV AGSTEYQKVAAESTIGSFDAN HQRVRAGGNTSGREYSKSVHSLVASETIGSFDAN HQRVRAGGNTSGREYSKSVHSLVASETIGSFDAN HQRVRAGGNTSGREYSKSVHSLVASETPRS HNGNELVESNEKGESSIYISDLNDKRKQKPA ENPCEGGSKNRNTEDSVQGSYFRAKPQKSPY GEGSGEFKKDGEFSVPSSNVEEVQKARAKK YIEHRISNETSVHSLPFDGVTGFPRGMLYECQ ECGECFAHSSDLTEHQKHDREKFSGSNYE WSVNSLAPTDPGTSYAQEQVAKQARNCK DFRQFFATSEDLNTNOKIVTOQEKSHGEBSQGE NTDGEETHSEETHGOFTIEDPVGKAPA SRCLUDSREYTHSVHTHSISESVQRDYTGEP KDPPDKYTECEDCGLGFVDLTDLTDHQYA SRCLUDSREYTHSVHTHSISESVQRDYTGEP SRCLUDSREYTHSVHTHSISESVQRDYTGEP SRCLUDSREYTHSVHTHSISESVQRDYTGEP SRCLUDSREYTHSVHTHSISESVQRDYTGEP SRCLUDGGFHISTSVHTHGDLLEQS QMAEAHIFOLALTEFQRSTEEHLFFCAVC ESFNYPAELADHVTHKNEPTYCTSSYTHTS FTITEPLKGAIPTYECKDOCKSFHISTVLTKHKE LHLEEEGEDBAAAAAAAQEVENVHVPQ VVLRIQGLNVEAAEFEVEAAEPEVEAAEPE EAAEPNGBAGGPOGEAAEFICAGONGBA QPRGDADEPDGVGIEDPEGEGDDDICQFOKAEPE GDADEPDGVGIEDPEGEGDDDICQFOKAEPE GDADEPDGVGIEDPEGEGDDDICQFOKAEPE GDADEPDGVGIEDPEGEGDDDICQFOKAEPE GDADEPDGVGIEDPEGEGDDDICQFOKAEPE GDADEPDGVGIEDPEGEGDDDICKRTSUSSAA SRRSSSDNRNFLERRAGLSSAAVQTRIGNSAV SRCCUARDARSTTEFTDFLQVALVRNRNQYFKM VCPAPRILARQLCXHILDPSCVALVRN VCPAPRILARQLCXHILDPSCVALVRN VCPAPRILARQLCXHILDPSCVALVRN VCPAPRILARQLCXHILDRSSVAA VCPAPRILARQLCXHILDRSSVAA SRCCUARDARSTTERRAGHLEGHLDFTKAALTH SRCCUARDARSTTARAGLEGHLDFTKAALTH SRCCUARDARSTTARAGLEGHLDFTKAALTH	1		ł				
GYEKSVIHSGPTESOKSHITTRPLESDEDEKA FITISSNPYENGKIPTKENVYPASSYERSVIRISL ASVEAGKSHSVAGPSKYRVMAESTIGSPANN HQRVRAGGNTSGGRYSKSVIRISLVASSPPRS HNGNEL VESNEKGESSIYISDLNDKRQKPAR ENPCEGGSKERNYEDSVIQSVFRAHPQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAHQKKSVP GEGSGEFKKDGETSVPQSSTRAPKOK DFPGFATSEDLNTNQKIVDGEKSIGEESQGE NTDGEETHSEETHGGETEDPVIQGSDMEPPQ KDDPDRIVIECEDCGLGFFVDLTDLTDHQKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFIHSSFLEHGHDDL 1SG RCLLCGGGFIHSALNEHMREDDLLEGS QMAEEAIPGLATTEPGRSQTEERIFFCAVCG ESFVNPALADNIVTHKNEHEMDDLLEGS QMAEEAIPGLATTEPGRSQTEERIFFCAVCG ESFVNPALADNIVTHKNEHYCOSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAAAGVEVYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAAAAAGVEVYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			1				
FTISSNPYENQKIPTKENVYEAKSYERSIQSFDANN ASVEAQKSHSVAGPSKPY VAGESTIQSFDANN HQRVRAGGINTSGGREYSKSVIHSLVASKPPRS HNGKELVESNEKGESSIYISDINDKROKEPAR ENPCEGGSKRAYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGESSVINSLDNOKROKEPAR ENPCEGGSKRAYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAKKK YEHRSNETSVHSLPFGCGTFRPROMLYECQ ECGECFAHSSDLTEHQKHIDERKPSGSRNYE WSVRSLAFTPDGTTSVAQESYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGESQGE NTDGETHSEETHQGTEIDFYQGSDMEDPQ KDPDDKIYECEDCGLGFVDLTDLTDHQKYH SRKCLVDSREYTHSVHITHSISEYQRDYTGQE LYECPKCGESFHISSFLFEHQRIHEQDQLYSM KGCLDGFRHSALNEHMRLHREDDLLEQS QMAEEAIFGLALTEFQRSQTEEHLFECAVCG ESPVIPABLADHVTVHKNEPYEVGSSYTHTIS FLIEFLKGAIFFYECKDCGKSFHISTVLTKKE LHLEEBEEDEAAAAAAAAAQSVEANIVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEVE EAAEPNGEAEGPDGEAAEFPGEAAEPEV EAAEPNGEAEGPDGEAAEFPGEAAEPEV GDADEPDGVGEDPGEGEDGQVEEPYYDC HECTETTTSSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANNQAGKEAPE GDADEPDGVGEDPGEGEDGQVEEPYYDC HECTETTSSTAFSHLKTHASMIFEPANAFG ECSGYIERASTSTGGANNQAGKVEANVHVPQ POSSEVPKLNVTVQDQEEHEXCOVGOQ LFNDHLSLARHONTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLERRAGLSSAAVQTRIGNSA ASRRSPAARPVVAPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYHV PPCSSEVPKLNVTVQDQEEVEKGFQCALSLQ HLRHWDPQEVTLQLFTDGTINKLIGCVVGN TIMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGBALDPKH VCNPAFFLLIARQLAKHAHRAWFMSSNL UKLMGKYSLIPTGPADEDINKRFLSDIPSSOI LQEEMTWMKERLSNI GSPVLONDALSHIFF GAVSVDVDYSLYPDBELQSQWLRAYLAYK EFRGGTEVTEKEVELLFIQVNQFALASHIFF GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVYP GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVYP GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVYP GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVYP GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVYP GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KERRICHTWARGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDBERVISTERGQHLGGGGF FFETSAKDININYKQTFERCI YDILCICKMSSSUK EKRILQIWDTAGGERYRTITYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDBERVISTERGQHLGGGGF FFETSAKDININYKQTFERCI YDILCICKMSSSUKS EFRSAKDINDYKQTERCI YDILCICKMSSSUKS EFRSAKDININYKQTFERCI YDILCICKMSSSUKS TO PAPPINSKRVDDYVTILLHISSKGHISPAKUTS	!						GVERGAINGCOUTEGONGUTTED DE CONCOCA A
ASVEAÇKSHISVAGPSKPKVMAESTIGSEDAN HQRVRAGGINTSGGRYSKSVINSLVASKPPRS HNGNEL VESNEKGESSIYISLINDKRÇKPAR ENPCEGGSKRRIVEDSVQSVPRAKPQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAKÇK YIEHRSNETSVHISLYFGCGTFRPRGMI, PECQ ECGECFAHSSDLTEHQKHIDERKPSGSRNYE WSVRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIVDEKSFIGESQGGE NTDGETHSETHIGGETIEDPVIQGSDMEPDQ KDPDDKIYVECEDCAGLGFVDLTDLTDHOKYH SRKCLVDSREYTHSVHITHSISEVQRDYTGEQ LYECPKCGESFHISSALFEHQRIHEDQLYSM KGCDDGFALLPMKPRRNRAARENALAGSA IRCLLCQGFHISSALNEHMRLHREDDLLEQS QMAEAHIPGLATTEPQRSPGTERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGALPFYECKDCGKSFHISTVLTKHKE LHLEEEEEDEAAAAAAAQEVEANVHVPQ VVLRIGGLNVEAAPEVEVAABPEVAAAPEV EAAEPHGEAGEPDGEAAPPIGEAGPNGEAE QPNGDADEPDGVGIEDEEGEQEQVEEPYYDC HECTETITISTSTAFSEHLKHASMIFEPAANFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHITG  1342 2692 A 10393 2 1350 GRPRSSSSINSNFIVERRAGISSAAVQTRIGNSA ASRRSPARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVAVVAVAVAMANYHVP PPGSPEVPKLINVTVQDQEERRCEAGLSLLQ HLRPHWDPQEVTLQLFTGGTINKLIGCYVGN TMEDVVLVRIYGNKTELL VDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAFFELLARQLAKHAHAHNGWPKSNL ULKMGKYFSLIPTGFADEDINKRFLSDPSSQI LQEEMTWMKELISNLGSFVLCENDLLCKNII VNEKQGDVGPIDYEYSGYNYLAYDIGNHFNE PAGYSDVDYSL VPDRELQSOWLRAYLEAYX EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALDQAKYSTEEPPGLYATVRGAMGFIN NYREKQGDVGPIDYEYSGYNYLAYDIGNHFNE PAGYSDVDYSL VPDRELGOSOWLRAYLEAYX EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALDQAKYSTEEPPGLYATVRGAMGFIN KERKILQIWDTAGQERYRITTAYRGAMGFI LMYDITNEESFNAVQDWSTQKTYSWDNAQ VILVGNKCDMEDERVSTERGQHLGGIFG FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSSLGF FFETSAKDNINVKQTTERLYDTLGCKMSSLGF FFETSAKDNINVKQTTERLY	]					j	ETICONDVENIOVIDEVENIOVE A VOVEDOVINIO
HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS HNGNEL VESNEGGESSIYISDLNDKROKDAPAR ENPCEGGSKRANYEDSVIQSVPRAKPQKSVP GEGSGFEKKOGESSVYSDLNDKROKDAPAR ENPCEGGSKRRAYEDSVIQSVPRAKPQKSVP GEGSGFEKKOGESVYSDLANAKKK YJEHRSNETSVIHSLPPGEQTFRPROMLYECQ ECGECFAHSSDLTEHQKHIDERKPSGSRNYE WSVRSLAFTPDGTTSVAGESVAGARAKKK DFRQFFATSEDLNTNQKIYDQEKSHGEBSQGE NTDGETHSEETHQGTIEDPVIQGSSMEDPQ KDPPDDKIYECEDCGLGFVDLTDLTDHOKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFIHSSILFEHQRIHEQDQLYSM KGCDGFIALIPMFRPRNARARPALAGSA IRCLLCGQGFHISSALNEHMRH HREDDLLEOS QMAEAHPGLALTEPGRSPGTEBLFECAVCG ESFVNPAELADHVTVHKNEPYEVGSSYTHTS FLTEPLKGAIFFYECKDCGKSFHISTVLTKKE LHLEEBEEDEAAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGVGIEDPEGEGDGAOGPNGEAE QPNGDADEPDGVGIEDPEGEGDGAOGPNGEAE QPNGDADEPDGVGIEDPEGEGDGAOGPNGEAE QPNGDADEPDGVGIEDPEGEGDGAOGPNGEAE GEGSTIERASTSTGGANQAGKECOVCOQ LFNDHLSLARHONTHITG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVVAPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMNYHV PPGSFEVPKLNVTVQQDEHERCEGGALSLLQ HLRPHWDPQEVTLQLFTDGTINKLIGCVVGN TMEDVVLVRIVGNKTELLYDGNEVKSFRVL QAHGCAPQLYCTFNNGLCYEFICGEALDPKH VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILLARGLAKHAHHANGWFKSNL VCNPAFFILARGLAKHAHHANGWFKSNL VCNPAFFILARGLAKHAHHANGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHATHANGWFKSNL VCNPAFFILARGLAKHATHANGWFKSNL VCNPAFFILARGLAKHAHTARNGWFKSNL VCNPAFFILARGLAKHAHTARNG						ļ	A SVE VONCHENA COCKDANA VECENOCED VD
HNGNEL/VENNEGESSIYISLINDKRQKPAR ENPCEGGSKRMYEDSVIQSVPRAKPQKSVPY GEGSGEFKKDGEFSVPSSNVREYQKARAKK YTEHRSNETSVIHSLPFGEQTFRPRGMLYECQ ECGECRAHSSOLTEHQKHIDREKPSGSRNYE WSVIRSLAFIDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLINTNQKTVDEKSFIGESGOGE NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ KDPPDKIYECEDCGLGFVDLTDLITDHOKVH SRKCLVDSREYTHSVHITHSISEYQRDYTGEQ LYECPKCGESTHISSFLFEHQDLLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQQGFHSSALNEHMRI.HREDDLLEQS QMAEEAIPGLALTEFQRSQTGETERLFECAVCG ESFVNFALIADHVTVHKNEPYEYGSSYTHTS FLITEPLKGALPFYECKDGKSFHSTVLTKHKE LHLEEEEEDEAAAAAAAQEVEANVHVPQ VVLRIQGLNVBAAFPEVEAABPEVAAPPEV EAAEPNGEAEGPDGEAAPIGEAGQPNGEAE QPNGADAEPPGAGIEDPEREPGAAPPE GDADEPPGVGIEDPEGEDQEIQVEEPYYDC HECTETTSTSTAFSEHLKTHASMIFEPFANAFG ECSSYIBRASTSTGGANQADEKTYFKCDVCGQ LFNDHLSLARHQNTHTIG  1342 2692 A 10393 2 1350 GRPRSSSDNRYNFLERRAGISSAAVQTRIGNSA ASRRSFARPPVPAPPALFRGRPGTGESTSLS APAVLVVAVVAVVAVAVAMAMNYHVY PPGSFEVPKLNVTVQDQEBERCREGALSLQ HLRPHWPPQEVTLQLFTDGITNKLIGGSYUGN TMEDVVLVRIJGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGBALDPKH VCNPAFFILLARQLAKHAHAHANGWPKSNL ULKMGKYPSLIPTGFADEDINKRFLSDPSQI LQEEMTWMKKELSNLGSSVVLCHNDLLCKNII VNEKQGDVGTDYEYSGYNYLAYDIGNHFNE FAGYSDVDVSLYPDRELQSORLAYLEAYAY EFRGGTEVTEKEVEILFIQVNQFALASHFFW GLWALDQAKYSTIEFPFLGYAIVAFNQYFKM VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILLARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD,VRENELGRGALDPKH VCNPAFFILARD  1343 2693 A 10394 102 839 PEAGTSAVLAREKGHLFTMRHEAPMOMASA QDARYGKORSULET  PEAGTSAVLANDRAGERVISTERGQHLGEQLGFE	1						
ENPCEGGSKNRNYEDSVIQSVPRAKKK GEGSGSEFKKDGESVYSSNVREYQKARAKKK YEHRSNETSVHISLPFGEQTFRPRGMLYECQ ECGECFAHSSDLTEHQKHIDDEKFRSGRNYE WSVIRSLAPTDPQTSY AQEQY AKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGESQGGE NTTGEETHISEETHIGQETIEDPVIQGSDMEDPQ KDPPDKIYECEDCGLGFVDLTD1THQKVH SRKCLVDSREYTHSWHTHSISEYQRDYTGEQ LYECPKCGESFHISSFLFEHQRIHEQDQLYSM KGCDDGFHALLPMKPRRNRAAERNALAGSA IRCLLGQGGFHSSALNEHMLHHEDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFCAVCG ESFVNPAELADHVTVIKHNEY LHEEEEDBAAAAAAAQVEANNYHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAQPEGAAGPROEAE QPNGDADEPDGVGIEDPEEGEAQQPNGEAE QPNGDADEPDGVGIEDPEEGEAQEPROEA QPNGDADEPDGVGIEDPEEGEAQQPNGEAE QPNGDADEPDGVGIEDPEEGEAQQPNGEAE PEGETETTSSTAFSEHLKHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLABHQNTHTI  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSAARPPVPAPAPALPRGRPGTGSTSLS APAVLVVAVAVVAVVSAVAWAMANYHV PPGSPEVPKLNVTVQDQEEHCRGEALSLQ HLRHWDPGVTUQLTDGTINKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLYEFIQGBALDPKH VCNPAIFRLIARQLAKHAHAHNGWIPKSNL WLKMGKYFSLIPTGADEDINKRFLSDIPSSQI LQEBMTWMKELISNLGSFVVLCHNDLLCKNII YNEKQGDVQFIDVEYSGYNYLATDIGNHFINE FAGYSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIQAKYSTEFDFLGSWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNGFALSHFFW GLWALIGAKAGAG UNDAGAGAG UNDAGAGAG UNDAGAGAGAG DPDWNSRVDDFWNLHILISSKGHISFAADTS			ļ				
GEGSGEFKKDGEFSVPSSNVREYQKARAKKK YTEHRSNETSVHISLPPGGITRPRGMILYECQ ECGECTAHSDLTEHQKHIDREKRPGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFROFFATSEDLINTOKYNYDQEKSHGEESQGE NTDGETHSEPHIQGETIEDPVIQGSMEDPQ KDPPDKIYDCEDCGIGGFVDLTDLTDHOKVH SRKCLVDSREYTHSVHTTHSISSYQRDYTGEQ LYECPKCGESFHSSELFEHQRITEQDLYSM KGCDDGFIALLPMKYRRNAAERINPALAGSA RCLLCGQGFHISSALNEHMRI.HERDDLLEQS QMAEFALIPGLALTEFORSOTERI.FECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLITEPLK.GAHFYYECKDGSFHSTVI.TKHKE LHLEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEFEVYDC ESAAFNGEAGFDGEAAAEPROEAGOPNGEAE QPNGDADEPDGAGIEDPEERAEFERGKAEFER GDADEPDGVGIEDPEEGGEFQKEEPER GDADEPDGVGIEDPEEGGEQVEFYYDC HECTETTSTSTAFSEHLKTHASMIHEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCOQ LFNDH.SLARHQNIHTIG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLERRAGLSSAAVQTRIGNSA APAVLVVAVAVVVVVSAVAWAMANYHVP PPGSFEYPKLNTVQDQEBHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLDVDDEEKSFRVL QAHGCAPQLYCTFNNGLCYFFQGEALDPKH VCNPAIFRLIARQLAKHHAHHAHDHKRPLSDFSOG LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLGIGNSHPISSOG LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLGIGNSHPISSOG LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLGHASHFPW GLWAHGLAKHHAHHAHDHKRPLSDFSOG LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYESGYNYLGHASHFPW GLWAHGLAKYSTEETDFLGYAIVRFNQYFKM KPEVTALKVPE  BAGSSDDANGTLERFAGGGISSAAVGTRIGNSA CDARYGCRDSSDQNFDYMFKLLLIGNSSVOK TSFLFRYADDSFTSAFVSTVGIDFKKTVFKN EKRIKLQWDTAGQERRYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTOJKTYSWDNAQ VILVGNKCOMBERKVISTERGOHLGGGF FFETSAKDNDVKQTFERLVDIICDKMSESLET DPAITAAKQNITELESKGHLGFAKDTS  DPAFTMAKQHTHLSSKGHISFAKDTS  DPAFTMAKQHTHLSSKGHISFAKDTS			[				
VIEHRSNETSVHISLPFGEQTERREMILYECQ   ECGECFARSDL TEHQKHEDEKPSGSRNYE   WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK   DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE   NTDGEETHSEETHGQETTEDPVIQGSDMEDPQ   KDDPDKIYECEDCGLGFVDLTDLTDHQKVH   SRKCLVDSRSYTHSVHTNISISEYQRDYTGEQ   LYECPKCGESFHISSELFEHQRIHEQDLYSM   KGCDDGFIALIPMKPRRAFRRAFENPALAGSA   IRCLLCQQGFHSSALNEHMRLHREDDLLEQS   QMAEEAIIPGLALTEFQRSQTERTHECAVCG   ESFVNPAELADHVTVIKHKEPYEYGSSYTHTS   FLTEPLKGAIFTYECKDCGKSFHISTVLTKHKE   LHLEEEEEDBAAAAAAAQVEANNHVPQ   VVLRIQGLNVEAAEPEVEAAEPEV   VVLRIQGLNVEAAEPEVEAAEPEV   EAAEPNGEAGEOPDGEAAAQVEANNHVPQ   VVLRIQGLNVEAAEPEVEAAEPEV   COLDEPPOVGEIDPEEGEDQEIQVEPTYDC   HECTETTSSTAFSEHLKHASMIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LYPDHISLARHQNTHTO   HECTETTSSTAFSEHLKHASMIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LYPDHISLARHQNTHTO   HECTETTSSTAFSEHLKHASMIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LYPDHISLARHQNTHTO   HECTETTSSTAFSEHLKHASMIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LAPINDHISLARHQNTHTO   HECTETTSSTAFSHLKTHASMIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LAPINDHISLARHQNTHTO   HECTETTSSTAFSHLKTHASMIFEPANAFG   ECSGYIERASTSTGANQADEKYFKCDVCGQ   LAPINDHISLARHQNTHTO   HECTETTSSTAFSHLKTHASMIFEPANAFG   ECSGYIERASTSTGANQADEKYFKCDVCGQ   LAPINDHISLARHQNTHTO   HECTETTSSTAFSHLKTHASMIFEPANAFG   ECSGYIERASTSTGANQADEKYFKCDVCGQ   LQEMTVMKELISHLGSFVLCHNDLLCKNI   YNEKQGDVQFIDVEYSGYVLCHNDLLCKNI   YNEKQGDVQFIDVEYSGYNYLADIGNHENE   FAGYSDVDYSLYPDRELQSQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWLRAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEVEILFQQWURAYLEAYK   EFKGFGTEVTEKEV							
ECGECFAHSSILTEHQKIHDRERPSGSRNYE   WSVRSLAPTDPQTSYAQEQYAKEQARNECK   DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE   NTDGETHISEETHQGTIEDPVIQGSDMEDDQ   KDPPDKIYECEDGCIGFVDLTDLTDHQKVH   SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ   LYECPKCGSFIHISSFLFHQRHEQDQLYSM   KGCDDGFIALLPMKPRRNRAAERNPALAGSA   IRCLLCQGFIHSSALNEHMRLHREDDLLEQS   QMAEAIIPGLALTEFQRSQTEERLFECAVCG   ESFVNPABLADHVIVHKNEPYYGSSYTHTS   FLITEJLKGAIPFYECKDCGKSFIHISTVLTKHKE   LHLEEBEEDEAAAAAAAAQEVEANNHVPQ   VVLRIQGLNVEAAFPEVAAFPEVBAAEPEV   EAAEPNGEAEGPDGGAGEDPEGCAEPE   GDADEPDGVGIEDPEFVYDC   HECTETTTSSTAPSEHLKTHASMIFFPANAFG   ECGYUREASTSTGGANQADEKYFKCDVCGQ   LFNDHLSLARHONTHTG   GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA   ASRRSPAARPPVAPPALPRGRPGTEGSTSLS   AAPAVLVVAVAVVVVVVSAVAWMANNYHW   PFGSFEVPKLNVTVQDEPHRCREGASLSLQ   HLRPHWDPQEVTLQLFTDGTITNKLIGCYVGN   TMEDVVLVRIYGNKTELLVDRIDEEVKSFVL   QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH   VCNPAIFRLIARQLAKHAHAHANGWIFKSNL   WLKMGKYFSLIPTGADEDNKRFLSDIPSSOJ   LQEEMTWMKELSNLGSPVVLCHNDLLCKNII   YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE   FAGVSDVDYSLYPDRELQSQWLRAYLEAYK   EFKGFGTEVTEKEVELLFIQWNQFALASHFFW   GLWALLQAKYSTLEPFDLGYATVRNQYFKM   KPEVTALKVPE   FAGVSDVDYSLYPDRELQSQWLRAYLEAYK   EFKGFGTEVTEKEVELLFIQWNGALASHFFW   GLWALLQAKYSTLEPFDLGYATVRNQYFKM   KPEVTALKVPE   FETSAKDNINVKQTFRLVTIKTSSKGHLETNAKTSUDKY   FFFTSAKDNINVKQTFRLVTITAYYRGAMGFI   LMYDITNEESFNAVQDWSTGKTYSWDNAQ   VLVGNKCDMEDEEVTSTERGQHLGGGLGFE   FFFTSAKDNINVKQTFERLVDILCDKMSESLET   DPAITAAKQNTRUKCTPPPQFRVACAC   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DPAITAAKQNTRUKCTPPPQFRVACAC   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DPAITAAKQNTRUKCTPPPQFRVACAC   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLISSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKGHLESPAKDTS   DAPFWNSRVDDDFVTILLIHLSSKG			1		}	ļ	
WSVIRSLAPIDPQTSYAQEQYAKEQARNKCK. DPROFFATSEDLNTNQKIYQEKSHGEESQGE NTDGETHSEETHQETIEDPVIQGSDMEDPQ KDDPDKIYECEDCGLGFVDLTDLTDHOKVH SRKCLVDSREYTHSVHITHSISEYQRDYTGEQ LVECPKCGESFHISSELFEHQRHIEQDOLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLICQGFHISSALNEHMRIHREDDLLEQS QMAEEAIIPGLALTEPQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHISTVLTKHKE LHLEEEEEDEAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAQQPNGEAE QPNGGDADEPDGAGIEDPEEGEDQEIQVEPYYDC HECTETTSTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYPKCDVCGQ LFNDHLSLARHQNTHTIG  1342 2692 A 10393 2 1350 GPRSSSDNNFFLERRGGLSSAAVQTRIGNSA ASRSPAARPPVAPPAPALPRGRRGTEGSTSLS APAVLVVAVVVVVVSAVAWAMANYHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFFLIARQLAKHAHAHAHNGWPKSNL WLKMGKYTSLIPTGPADEDINKRILSDIFSSQI LQEEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKQGDVOPIDYETSGTPNLAYDIGNHFNE PAGVSDVDYSLYPDRELGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEEPDFLGYAIVRNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PPAGTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKKRTVFNK KERIKLQIWDTAGGERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTGIKKKITVFNK EKRIKLQIWDTAGGERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTGIKKKITVTSWDNAQ VILVONKCDMEDERVISTERQQHLGEQLGFE FFETSAKDNINVKQTFERLVDILICKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLISSKGHIEPPAKDTS							
DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGEETHSEETHGGETIGEPVIQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEG LYECFKCGESFHISSETHEGRHEGQDLYSM KGCDDGFIALLPMKPRRNRAARNPALAGSA RICLLCGGFHISSALNEHMRIHREDDLEQS QMAEEAIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCKSFHISTVLTKHKE LHLEEEEDEAAAAAAAAAAQEVEANNHVPQ VVLRIQGLNVEAAFEVEAAEPEVEAAEPE EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGACPEN GADEPDGVGIEDPEGAGIEDPEERAEEPEGACPEN GADEPDGVGIEDPEGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGAGIEDPEGVEEPYYDC HECTETTISTAFSEHLKTHASMIIFEPANAFG ECSGYURASTSTGGANQADEKYFKCDVGQ LFNDHLSLARIQNTHTG  1342 2692 A 10393 2 1350 GFRFSSSDNRNFIRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMNYTHY PPGSFEVPKLNVTVQDGEHKCREGALSLLQ HLRPHWDPQGVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAHAHNGWIFKSNL WLKMGKYTSLIPTGADEDINKRILGDINSPISOI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YMEKQGOVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILIFIQVNQFALASHIFFW GLWALLQAKYSTEIPTGADEDINKRYLSDIPSSOI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YMEKQGOVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILIFIQVNGFALASHIFFW GLWALLQAKYSTEIPTFPLGYATVRNGYYKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTIMRHEAPMOMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTARVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESSNAVQDWSTQHKTYSWDNAQ VLVGNKCDMEDERVISTERGQHLGGGLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNITRLKETPPPPCPNACAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDPVTVILLILLSSKGHIESPAKDTS	1		ĺ			1	
NTDGETHSETHGQETIEDPVIQGSDMEDPQ KDDPDDKIYECEOGLGFVDLTDLDHQKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLIPMRRNAAGRNPALAGSA IRCLLCGQGFHSSALNEHMRLHREDDLLEQS QMAEEAIJFGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVIVHKNEPYEYGSSYTHTS FLTEPLKGAIFFYECKDCGKSFHSTVLTKHKE LHLEGEEDEAAAAAAAAAQEVEANNHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGGGEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDDQIQVEEPYYDC HECTETTTSSTAFSEHLKTHASMIIFPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNHTHG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LPNDHLSLARHQNHTHY PROSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTIGGTINKLIGCYYGN TMEDVVLVRIYGKNELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYFEIQGEALDPKH VCNPAFRLIARQLAKHAHAHANGWFKSNL WLKMGKYFSLPTTGADEDINKRFLSDIPSSQI LQEEMTWMKELSNLGSFVVLCHNDLLCKNII YNEKQGDVQFTDVFYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSWLRAYLEAYK EFKGFTEVTEKEVELLFQNOPGALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMKKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITITAYYRGAMGFI LMYDITNEESFNAVQDWSTQKTYSWDNAQ VILVGNKCDMEDBERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPPNCAC			ļ .				
KDDPDKIYPECEDCĞLGFVDLTDLTDHOKVTI SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFHESFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNKAAERNPALAGSA IRCILCQGCFHSSALNEHMRLHREDDLLEGS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEYPEYGSSYTHTS FLTEPLKGAIPFYEXCBCGKSFHISTVLTKHKE LHLEEEEDBAAAAAAAAAQEVEANVHVPQ VVLRIGGLNVEAABPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPPDGGEIDPEEGEDGEJOVEPFYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LNNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSAAPPVPAPLPRGRPGTGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSFEVFKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGTITNKLIGCYVGN TMEDVVLVAIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAFFRLIARQLAKHAHAHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRLSGVYSCN TMEDVVLVAIVANVAVVVVVSAVAWAMANYIHV PRGSPOVFKLNTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGTITNKLIGCYVGN TMEDVVLVAIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAFFRLIARQLAKHHAHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRLSGVYNCHDBLCKNII VNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTETTEKEVEULFIQVNQFALASHFFW GLWALIQAKYSTEFDFLGYAIVFRNQYFKM KPEVTALKVPE GLWALIQAKYSTEFDFLGYAIVFRNQYFKM KPEVTALKVPE GLWALIQAKYSTEFDFLGYAIVFRNQMSAA QDARYGGKDSSDQNFDYMKLLIIGNSSVGK TSFLFRYADDSTSAVSTVGDTPKKYTYFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITINEESPNAVQDWSTJCKTYSWDNAQ VILVONKCDMEDBERDERVISTERGOHLGEQLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNITLKETPPPPPNCAC				ļ			
SRRCL/DSREYTHSV/HTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLIMKPRRNRAAERNPALAGSA IRCLLCGQGFHSSALNEHMRLHREDDLLEQS QMAEAEJIPGIALTPGRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAJPFYECKDCGKSFHISTVLTKHKE LHLEEEEDEAAAAAAAAAQEVEAANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGAGQPRIGAEA QPNGDADEPDGAGEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGGEDQEJQVEEPYYDC HECTETTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GGPRSSSDNRFILRERAGLSSAAVQTRIGNSA ASRRSPAARPVPAPAPPALPRGRPGTEGSTSLS AAVAUVVAVVAVVAVVAVVAVVAVAWAMANYIHV PPGSFEVFKLNNTVQDQEEHRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TIMEDVVLVNIYGNKTELL VDRDEEVKSFRVL QAHGCAPQLYCTFNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRPLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDVEYSGYNYLAYDIGNHENE FAGVSDVDYSL YPDELQSQWLRAYLEAYK EFKGFGTEVTEKEVELFIQVNQFALASHFFW GLWALIQAKYSTTEPDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMKLLIIGNSSVGK TSFLFRYADDSSTSAFVSTVGIDFKKVTYFKN KPEVTALKVPE GLWALIQAKYSTTETDFLGYAIVRFNQYFKM KPEVTALKVPE LWYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDBERVISTERGQHLGEGLGFE FFETSAKDNINVKQIFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFYNTLHLISSKGHISPAKDTS		İ	ļ				
LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALIPMKPRNRAAERNPALAGSA IRCLLCQQGFIHSSALNEHMRIHREDDLLEQS QMAEEAIIPGLAATTEPGPRSQTEERLFECAVCG ESFVPAELADHVTVHKNEPFYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVBAAEPEVEAAEPFEVAEAPFEV EAAEPNGBAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGGAGEOPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDDEIQVEEPPYTOC HECTETTISTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGTINKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLVCTFNNGLCYEFIGGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKESILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYETYSGYNYLAYDIGNIFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVYTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLIRYADDSFTSAVYSTVGBDFKVTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDDEVNSTORGGOHLGFGLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC			ł				
KGCDDGTALLPMKPRRNRAAERNPÄLAGSA   RCLLCGQGFHSSALNEHMRLHREDDLLEQS   QMAEEAIIGELATTEPQRSQTEERLFECAVCG   ESFVNPAELADHVTVHKNEPYEYGSSYTHTS   FLITEPLKGAIPFECKDCGKSFHSTVLTKHKE   LHLEEEEDEAAAAAAAAAQEVEANVHVPQ   VVLRIQGLNVEAAEPEVEAAEPEV   EAAEPNGEAGGPGGEAAEPIGEAGQPNGEAE   QPNGDADEPDGAGIEDPEEGEDQEIQVEAAEPEV   GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC   HECTETTSTSTAFSEHLKTHASMIIFEPANAFG   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LFNDHLSLARHQNTHTIO   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LFNDHLSLARHQNTHTIO   ECSGYIERASTSTGGANQADEKYFKCDVCGQ   LFNDHLSLARHQNTHTIO   PPGSFEVPKLNVTVQDQEEHRCREGALSLQ   HLRPHWDPQEVTLQLFTDGTTNKLIGCYVGN   TMEDDVLVRYVQVSAVAWAMANYIHV   PPGSFEVPKLNVTVQDQEEHRCREGALSLQ   HLRPHWDPQEVTLQLFTDGTTNKLIGCYVGN   TMEDDVLVLRYDKTELLVDRDEEVKSFRVL   QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH   VCNPAIFRLIARQLAKHAHAHANGWIPKSNL   WLKMGKYFSLIPTGFADEDINKRFLSDPSSQI   LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII   YNEKQGDVQFIDYEYSGYNYLAYDIGNIFNE   FAGYSDVDYSLYPDRELQSQWLRAYLEAYK   EFKGFGTEVTEKEVEILFIQVNQFALASHFFW   GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM   KPEVTALKYPE   FAGYSDVDSLYPDRELQSQWLRAYLEAYK   EFKGFGTEVTEKEVEILFIQVNQFALASHFFW   GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM   KPEVTALKYPE   KERIKLQIWDTAGQGEYRTITTAYYRGAMGFI   LMYDITNEESFNAVQDWSTQIKTYSWDNAQ   VILVGNKCDMEDERVISTERGQHLGEQLGFE   FFETSAKDNINVKQTFERLVDIICDKMSESLET   DPATTAAKQNTRLKETPPPPQPNCAC   1344   2694   A 10395   2 4136   DRPPWNSKUDDFVTNLHLISKGHISFAKDTS	1		1	'			
IRCLLGQGFIHSSALNEHMRLHREDDLLEQS QMAEAHIGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVITVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETFISSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARRQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSILLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAIHAHNGWIPKSNL WLKMGKYFSLLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN KPEVTALKVPE EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGGHLGFGLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNITILKETPPPPPQNCAC  1344 2694 A 10395 2 4136 DRPWNSRVDDFVTNLIHLSSKGHISPAKDTS			İ				
QMAEEAIPGLALTEFQRSQTEERLFECAVCG ESFVAELADHVIVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNEAAEPEVEAAEPEVEAAEPEVE EAAEPNGEAEGPDGEAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPERGEEPPGKAEEPE GDADEPDGVEEPPEGGEDGEIQVEEPYDC HECTETFTSSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAMAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRHWDPQEVTLQLTDGITNKLIGCYYGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYPSLIPTGFADEDINKEFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKGQDVQFIDYSGTYNLCHYDGINKHFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKGQDVQFIDYSGTYNLCHYFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLGSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLQSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLQSQWLRAYLEAYK FFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALIQAKYSTIEPDFLQSQWLRAYLEA FFTSAKDNINVKQTFERLYDDICCKMSESLET DPAITAAKQDITRILKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPFWNSRVDDFVYNLLHLSSKGHISPAKDTS	1		ł				
ESFVNPAELADHYTYHKNEPYEYGSSYTHTIS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAGEPDGEAGPNGEAE QPNGDADEPDGAGIEDPEEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCVYGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFELIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVGPUDVEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYFE FKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSDQNFDYMFKLLIIGNSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESANAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLVDIICDKMSESLET DPAITAAKQNINVKQTFERLYDDIICDKMSESLET DPAITAAKQNINVKQTFERLYDDIICDKMSESLET DPAITAAKQNINVKQTFERLYDDIICDKMSESLET DPAITAAKQNTRILKETPPPPQPNCAC	i I						IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS
FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAAFPEVEAAAPEVEAAAPEVE EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDOVGIEDPEBGEDDEIQVEEPTYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVVSAVAWAMANYIHV PPGSFEVFKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGADEDINKRFLSDIPSSQI LQEEMTWMEELSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK KPEVTALKVPE LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERQQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLSSKGHISPAKDTS	1					•	QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
LHLEEEEDBAAAAAAAAAQVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAAGPPGEAAGPPCAAEPEV EAAEPNGEAAGPPGEAAGPPGEAAGPPG GDADEPDGAGIEDPEGEDOEIOVEEPYYDC HECTETIFTSSTAFSEHLKTHASMIFPEANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHONTHTG GCSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHONTHTG ASRRSSDNRFFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFIARQLAKHAHHAHNGWPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFJQVNOFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE DAARGAGGRYRTITTAYYRGAMGFI LMYDITNESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLSSKGHISPAKDTS							
VVLRIQGLNVEAAEPEVEAAEPEVE EAAEPROVEAAEPEV EAAEPNGEAEGPOGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEEGEDEPEGKAEEPE GDADEPDGVGIEDPEEGEDPEGVEEPYYDC HECTETTSTAFSEHLKTHASMIIFFEPANAFG ECSYUREASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNPLRERAGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFIDGITNKLIGCYVGN TMEDVVLVRIYGNKTELL VDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHAHNGWIPKSNL WLKMGKYPSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPOPNCAC	1		į				
EAAEPNGEAEGPDGEAAEPIGEAGOPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLQ HLRPHWDPQEVILQLFIDGITINKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFIRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTELKETPPPPQPNCAC							
QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSAARPPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNOFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAHTAAKQNTRLKETPPPPPQNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLSSKGHISPAKDTS			ĺ				
GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELLSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNOFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSTTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAHTAAKQNITRLKETPPPPPQNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				′			
HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCQQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPSSSDNRNPLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVYKLNVTVQDQEEHRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELLSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1						
ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG  1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPVPPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAFFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNIL YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	ł			}			
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNTRLKETPPPPPOPNCAC	1			ì			
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGPTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFIDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTOFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				
ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							LFNDHLSLARHQNTHTG
APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1342	2692	A	10393	2	1350	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			!				
HIRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNIFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				APAVLVVAVAVVVVVVSAVAWAMANYIHV
HIRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNIFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			J	1			PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			]				
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	Ì		1				QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1	1	1	ļ			
YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1	ĺ	ĺ	1			
FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							YNEKOGDVOFIDYEYSGYNYLAYDIGNHFNE
EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	[		[	[			
GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				EFKGFGTEVTEKEVEIL FIOVNOFAL ASHFFW
KPEVTALKVPE  1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1			}	
1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1	ļ		ļ	
QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1343	2693	A	10394	102	839	
TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			١	.55,74		1 327	
EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	İ		l				TSELER VADDSETS A EVET VCIDER VINT VERNI
LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS		]	1			,	
VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				
FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		I				ON A CORCUME DED AGAINST A COLOR OF COL
DPAITAAKQNTRLKETPPPPQPNCAC  1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS		Ì	I				
1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1	ľ	{		ĺ	· '	
The state of the s	1344	2604	<u> </u>	10205	-	4126	
LQQRIFAEMSFVLHFYVRPSGHEGAASGHTR	1344	2094	[ A	כלכטו	2	4130	
		L	L	L	L	L	LUCKITALINIST VLHI Y VKPSUHEGAASGHTR

NO. of polito coide of the control o							
peptide cetotide sequence unence companie de la companie	SEQ ID		L	SEQ	Predicted		Amino acid sequence (A=Alanine C=Cysteine,
peptide cetotide sequence unence companie de la companie	NO: of		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence   Sequence			}	in		location	F=Phenylalanine, G=Glycine, H=Histidine,
uence   09/495   correspondi   a formation   a formation   acid residue   cold			Ī				
uence   914			l .	1			
amino acid residue of peptide sequence propried sequence propried sequence provide sequence s		uciice	1			,	
residue of peptide sequence    Y=Tyrosine, X=Usknown, **Siop odon, *possible nucloside detelor, \( \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \sip \) possible nucloside detelor, \( \sip \sip \) possible nucloside detelor, \( \sip \sip \sip \) possible nucloside detelor, \(	uence	}	l	914			
# peptide sequence   #-possible nucleotide deletion, *-possible nucleotide insertion   nucleotide insertion   nucleotide insertion   RKLQGKLPELGOVETELCYNNOWTAEALFS,   EFIKKLMWLFGGEVLLDDVARESVLLPGSN   DLLLEVGRENNFSTFSTNIVSVCRATGLGPV   DRVETTRRYRLSRAPPPSAPERALATALHAD   MTEOHPPHFIQSSPSEMPEPLNOPNINLGEGE   LALEKANQELGIALDSWDLDPYTREYREQUALATION   NEW PERSON		l	l				
Poptide   Popt			l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
nucleotide insertion  RKLOGKIPELGOVETELCYNNWTAEALES.  RETKKLMWLPGCPLLLDDVARESWLLPGSN DLLLEVQPRINNSTPTSTNIVSVCRATGLGPV DRVETTRRYRLSFAHPPSAEVALTGLGFV DRVETTRRYRLSFAHPPSAEVALATALH, DE MTEOHPHPIOISSPESMFENNEPINGHORING.GER LALEKANQELGLAL DSWDLDFYTKRPQELQI NPSTVEAPDIA_QSNESHBRINFEKGLHVD OKLVHSLFESIMSTOSSNPNNVLKFCDNSSN IQGKSVRTLRPEDFTRSRPQQQGLRHVVYT AETHNPPTGVCPSSATTGTGGRIRDVQCTG RGAAHVAGTGGVCFGNLHHIGTYNLPWEDLS; QYPGNFAPPLEVABEASNOASDYGAKFGEW LAGPARSLGQLPDGGREWWRINFSGGGS MEADHISKEAPERGBEVVKVGGDYNRIOVG GAASSYQVGDONTSDLDFTGAVQRGDFMERG KMRNVRACVEAFKGPFCSLHDGGAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSNALLLSRNFROFLTHVSARERCPA CFVGTTTGDRRIVLVDDRECFVRRNGGGDAF PTPPTTVDLELEWVLGKMPRKEFTLGKFP MLQPLALPFGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLGPTLADVA VVALSHEELIGAATALGEQFVKSLLDFKVAA RAVABALTINLVRAVTOTENVCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV, VDGGKDSLSMAARVGTTVRARGSLVSAY VCPDITATVTPDLKHPEGRGHLLYVALSPQQ HRLGGTALAQCFSQLGHPPDLDLPSN.VSA FSITQGLLKDRLCSGHDVSDGGLVTCLLEM AFAGNGQVDVPPVPVDVLSVLFAEEGGNAD LEVQEPDLAQVLKRYRDAGIHCLEHPTLSCHPPNLDLPSN.VSA FSITQGLLKDRLCSGHDVSDGGLVTCLLEM AFAGNGQLQVDVPVPVDVLSVLFAEEGGNAD LEVQEPDLAQVLKRYRDAGIHCLEHPTLSC LEVQEPDLAQVLKRYRDAGIHCLEHTING BMADFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDI FRGVAFVGGFSYAVALREGSSNDD EMADAFHLAGFEVVDVTMQDLCSGAIGGDLC GEKKHK TYLLAVQVBARAQVLTSSKCLITALHELKNI TYPSWLLDGFFRTLTQAEALDRAYQDTVN NVFFEVYKKRUPHFASREVYVHENPFFTLTTS FVQLDLFRGREDKFFTVKRLKAYEL TYGIDDLTGEFLIQREDDKFFTVKRLKAYEL TYGIDDLTGEFLIQREDKFTVKRLKAYEL TYGIDDLTGEFLIQREDKFTVKRLKAYEL TYGIDDLTGEFLIQRE	j i	ĺ	Ì	1	nentide	•	
RKLQGKIPELGOWETELCYNNWTAEALDS: ETKKLMWLFGGFULDDVARESULPGSN DLLLEVGPRINFSTFSTNIVSVCRATGLGPY DRVETTRRYRISPARSPARSULPGSN DLLLEVGPRINFSTFSTNIVSVCRATGLGPY DRVETTRRYRISPARSPARSULATATHE MTEGHPPHPIGSSPESMPEPLNOPINILGERG LALEKANDELGIALDSWDLDPYTRFQGLGV QKJVHSLFESMSTGDSSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNSS IQGKEVRILPGESSMPNVLKFCDNS ICGKESSMPNVLGESSMPNVLGENGVVKCGDEV LAGPASSIGLJJPGOQREWKFNRFSCGIGS MEADHSKEAPERPGEVVKVGGPVYRICOG GAASSVQVQGDNTSDLDFGAVQRGDFEME KMNRVRACYERSPNDTITHVSAVRECCA CFVGTTTCDRRIVLVDDRECFVRRNGQGDAP FTPPTTYDLLELEVALQGEVCVEPLOGAGGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQVQGNVTARAGGAGNG GAASSVQNVTARAGGAGNA RULVALAUTNITARAGGAGNG WAALLTNITARAGGAGNA RULVALAUTNITARAGGAGNG GAASSVQVQGNVTARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGNA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAGAA RULVALAUTNITARAGGAGAAA RULVALAUTNITARAGGAGAAA RULVALAUTNITARAGGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAATTIGGAGAGAAAAAAAA		i			,		
EETKKLMWI-FGCPLLLIDDVARESWIL-IPGSN DLLEUGRENNSTTSTINISVCKATGLGPV DRYETTREYRLSFAHPPSAEVEAIALATILIDE MTEOHPHPIOSSPESMPENLANPINLGEGR LALEKANQELGIALDSWDLDFYTKEPQELQI NPSTVEAPDIA_OSISEHSRIWFEKQUI-VOW QKLVHSLFESIMSTOESSNPNNVLKFCDNSS, IQKESVERLEPEDFTRSSPQQQGCHENVYF AETHNPPTGVCPSSCATTGTGGRRDVQCTG RQAHVVAGTGAVCFONLHIPGYNLPWEDLS QYPONPARPLEVAIEASNCASDYGNKFGEPV LAGPARSLGQLPDQGRREWKPMPSGGIOS MEADHISKEAPERGREWVXVGGPYNICIVG GAASSVQVQCONTSDLDFOAVQRGDFMRE KMRRVIRACVEARKGPPICSLHDGAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNROPLTHVAGAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNROPLTHVAGAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW GASYQSSAALLASRNACHECAGGNG NVLKELSDPAGAITYTSRPQLGDFTLNALEIW AVALSHEELIGAATALGEQFVKSLLDFKVAA RAAGNGCGGNGAALADCEAMVAVMAALGV VALSHEELIGAATALGEGVYKSLLDFKVAA RAAGNGCGGNGAALADCEAMVAVMAALGV VODGGKOSLSMAARVGTETVRARGSLINSAY, VCPDITATVTPDLKIPEGRGHLLYVALSPOQ HRIGGTALAQCFSQLGEHPPDLDLPFNLVAA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNGCQVVDFVYPVDVLSVLYALSPOQ HRIGGTALAQCFSQLGEHPPDLDLPFNLVAA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEW TSFQLDRLQAFRCVAEEERGLRERMGPSVV LEVQEPDLAQVLKRYRDAGIHCLELGHTGE AGPHAAVRVSVNGAVVLEPVOEERLALWE TSFQLDRLQAFRCVAEEERGLRERMGPSVV LEVQEPDLAQVLKRYRDAGIHCLELGHTGE AGPHAAVRSVNGAVVLEPVOERLANDH TSFQLDRLQAFRCVAEEERGRERMGPSVV LEVGFPASSPLAQLEARGRACHATHY ADDDGNPTGAVARWQVLEPVOERLANDH TSFQLDRLQAFRCVAEEERGRERMGPSVV LEVGFPASSPLAQLEARGARDATATHY NAMPYGMSVTLKYRDRAGGAVATFLFFGSAND NAMPYGMSVTLKYRDRAGGAVATFLFFGSAND NAMPYGMSVTLKYRDRAGGAVATFLFFGSAND NAMPYGMSVTLKYRDRAGGAVATFLFFGSAND NAMPYGMSVTLKYRDRAGGAPRAATTGSSPCDGRE RAAAHTNAWTLEGSC GVGCFFWAGTNAASRAGPRAATTGSSPCDGRE TVAMPYGMSVTLKYRDRAGGAPRAATTGSSPCDG TVAMPYGGSGAAMGASARLLRAVIMGAFGS GKGTVSSRITTHFELKHLSSGOLLROMLRG GKGTVSSRITTHFE			ļ		sequence		
DLLLEVGRINTSTITINIVS/CRATGLGRY DRVETTRAYRIS-RAHPPSAVEVALATALTHED MTEGHPHPHOSIS PESMPEPL NOPINIL-GERG LALEKANQELGI-LALDSWOLPYTKRYGELQ) NPSTVEARDLAQSNESHSRIWFEKQQLIAVO QKI_VHSLFESMSTOESSMPNVLKFCONSS, IQKEVRLRPEDPTRSSERQQQQLEHVYFT APTHNPPTGVCPTSCSATTGTTGGRIRDVQCTG RGAHVVAGTAGYCFGNLHIFGYNLPWEDLS; QYPGNFARPLEVALEANGSDYGNKFGERV LAGFARSLGQLPDGGGGREWIKFINFSGGIGS MEADHISKEAFEGMEVVKVGGFVYRIGVGG GAASSVQVGGDNTSTRROGLOFTINALEIW GASYQESNALLLRSPNEDFLTHVSARBECPA WILLESDPAGAUTYSRROGLOFTINALEIW GASYQESNALLLRSPNEDFLTHVSARBECPA CFVGTITGDRRIVLYDDRECPVRRNGQGDAP PTPPTTVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VALASHELIGAATALGEQFVSKLDPKVAA RLAVABALTNLVFALLTDLRDVKCSGNWM WAAKLPGGAALADACBAMVAWAALGV, VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKHPBGRGHLLYVALSPGQ HRLGGTALAQCFSQLGHPPDDLDEPDLLPENLVRA FSTGQLLKDRLLCSGHDVSGGILVTCLLEM AFAGNGGLOVDVPVPRVDVLSVLFAEEDENL LEVQEPDLAQVLRYRDAGLHCLLETTLE AGPHAMVRVSVNGAVVLEPVGLRALWEI TSFQLDRLQAEPRCVAEEERGLREMMGPS LEVGEPDLAQVLRYRDAGLHCLLGHTGE AGPHAMVRVSVNGAVVLEPVGLRALWEI TSFQLDRLQAEPRCVAEEERGLREMMGPS LEVGEPDLAQVLRYRDAGLHCLGHTGE AGPHAMVRVSVNGAVVLEPVGELRALWEI TSFQLDRLQAEPRCVAEEERGLARLEN SGRYESRWASVRVOPGPALMLRGMEGAVLL WOGGPHEDAABRGPBOQARGTLRHNL SGRYESRWASVRVOPGPALMLRGMEGAVLL VWSAHGGGVYALSPSPELQAQUEARGLALLWEI TSFQLDRLQAEPRCVAEEERGLAPLEY ADDDGNFTEQVPLNPNGSSGGVAGICSCDGE HLAVMPHPERAVRPQWAWRPPFDTLTTS PWLQLENNARNWTLEGSC UVRGGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYRIKKLIVYHLVIWLLL LYGEMEMPAAQQLYRHKKKRFTFORSAVS MYLUVLAVOVHAWQLYYSKKLLDSWFTS DEVKLRGGVYARSSPELQAQUEARGLAPLEY ADDDGNFTEQVPLNPNGSSGGVAGICSCDGE HLAVMPHERAVRYRWQWAWRPPFDTLTTS PWLQLENNARNWTLEGSC TSPCHWGGTYARSPELQAQUEARGLAPLEY ADDDGNFTEQVPLNPNGSSGGVAGICSCDGE HLAVMPHPERAVRRYWQWAWRPPFDTLTTS PWLQLENNARNWTLEGSC GVGGFWAGTMASRAGARAGRAAGTDGSDFQHR TVAHHYDGRAGHAPAQQLYRHKKLAYEL TGYSWLLDGFFRTLPQABALDRAYQDTVNN NYFFEVUKQULTARWHPASGCVANGENSTAPL TYGIDDLTGFRURPDDWTTLALHELKNI TQYSWLLDGFFRTLPQABALDRAYQDTVNN NYFFEVUKQULTARWHPASGCVANGEFTSTAFT TVGIDDLTGFRURPROGREDKETVIKRILAYED TYNFULTYYGKGULTENGETSTIKIWPYV AFLQFFRTLPGABALDRAYQDTVNN NYFFEV			İ		į		
DRVETTRRYRLSRAHPPSACVEALALATLING MTEOHPHPHOSESPSEMPENJOPNIN.GEOR LALEKANQELGLALDSWDLDFYTKEPQELQ) NPSTVEARTILAGNSEHSRIWFEKGGLIVOR QKLVHSLFESIMSTQESSNPNNVLKPCDNSSA (QKEVRRLPEDPTRRSEPQQQGLRHVVF) AETHNPFTGVCPFSCATTGTGGRREDQVCTG RGAHTVVAGTAGTGTGGTGROTHQVCTG RGAHTVVAGTAGTGTGGTGROTHQVCTG RGAHTVVAGTAGTGTGGTGROTHQVCTG RGAHTVVAGTAGTGTGGTGROTHQVCTG RGAHTVVAGTAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG		ſ	1	1		1	
MTECHTPHIQSTSTESMPETLAGPINITEGE     LALEKANQEL GLALDSWOLDFYTKEROGLEY     NPSTVEAPDLAQSNSEHSRHWFFKGOLIFYON     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     IQGEVVEFLRPEDPTEPSRPQQQGLRHVYFT     AETHNPTIOV.CPTSGATTGTGGRIRDVOCTG     RGHRVAGTAGYCFONLHIROFYNLPWEDLS!     QYPONFARELEVALEASNGASDYGNKFOEDS     QYPONFARELEVALEASNGASDYGNKFOEDS     AETHNPTIOV.CPTSGATTGTGGRIRDVOCTG     GARSIGLOJ.DPGQREWIKPIMFSGGIOS     MEADHISKEAPEPOMEVVKVGGPVYRICOV     GARSSVOVQGDNTSIDLOFGAVGROPEMSC     KMNRVIRACVEAPKGNPICSLHDQGAGGNG     GARSSVOVQGDNTSIDLOFGAVGROPEMSC     KMNRVIRACVEAPKGNPICSLHDQGAGGNG     NVLKELSDPAGAITYSRFQLOFTNALEIW     GAEVQESNALLLRSPNRDFLTHVSARERCPA     CFVGTTTGDRRIVLVDRECVERNY     GAEVQESNALLLRSPNRDFLTHVSARERCPA     CFVGTTTGDRRIVLVDRECVERNY     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLAVEALINLVFALVTOLRDVCCSGNWW     WAAKLPGEGALADACEAMVAVMAALGV     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VALSHEELIGAATALGEQPVSKLLDFKVAA     RLAVEALINLVFALVTOLRDVCCSGNWW     WAAKLPGEGALADACEAMVAVMAALGV     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDLSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     FRITGOLLKDRLLCSGHDWSDGLVTCLLEM     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGLGVTCLLEMT     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGLGVTCLLEMT     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGAICHCLEMTIC     AGAGLERRFRKRPDTTSLGVCNGCQLLALD     VWSAHGEGVYAFFECALARLEGSNGDD     EMADAFILAGFEV WDVMAWALARGGNAARVTFIF     RAGAELRFRKRRPDTTSLGVCNGCQLLALD     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT	1	ł	ŀ	İ			DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV
MTECHTPHIQSTSTESMPETLAGPINITEGE     LALEKANQEL GLALDSWOLDFYTKEROGLEY     NPSTVEAPDLAQSNSEHSRHWFFKGOLIFYON     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     OKLVHSL.FSIBATOGESSNPHYLYEODNSA     IQGEVVEFLRPEDPTEPSRPQQQGLRHVYFT     AETHNPTIOV.CPTSGATTGTGGRIRDVOCTG     RGHRVAGTAGYCFONLHIROFYNLPWEDLS!     QYPONFARELEVALEASNGASDYGNKFOEDS     QYPONFARELEVALEASNGASDYGNKFOEDS     AETHNPTIOV.CPTSGATTGTGGRIRDVOCTG     GARSIGLOJ.DPGQREWIKPIMFSGGIOS     MEADHISKEAPEPOMEVVKVGGPVYRICOV     GARSSVOVQGDNTSIDLOFGAVGROPEMSC     KMNRVIRACVEAPKGNPICSLHDQGAGGNG     GARSSVOVQGDNTSIDLOFGAVGROPEMSC     KMNRVIRACVEAPKGNPICSLHDQGAGGNG     NVLKELSDPAGAITYSRFQLOFTNALEIW     GAEVQESNALLLRSPNRDFLTHVSARERCPA     CFVGTTTGDRRIVLVDRECVERNY     GAEVQESNALLLRSPNRDFLTHVSARERCPA     CFVGTTTGDRRIVLVDRECVERNY     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLQFLALPPGLSVRIQALERVALARE     MLAVEALINLVFALVTOLRDVCCSGNWW     WAAKLPGEGALADACEAMVAVMAALGV     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VALSHEELIGAATALGEQPVSKLLDFKVAA     RLAVEALINLVFALVTOLRDVCCSGNWW     WAAKLPGEGALADACEAMVAVMAALGV     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDLSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     VDGGKDSLSMAARVGTETVRAPGSLUSSA     FRITGOLLKDRLLCSGHDWSDGLVTCLLEM     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGLGVTCLLEMT     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGLGVTCLLEMT     AFAGNCGLQVDVPVPRVDVLSVLFAEEDCL     LEVQEPDLAQVLKRYRDAGAICHCLEMTIC     AGAGLERRFRKRPDTTSLGVCNGCQLLALD     VWSAHGEGVYAFFECALARLEGSNGDD     EMADAFILAGFEV WDVMAWALARGGNAARVTFIF     RAGAELRFRKRRPDTTSLGVCNGCQLLALD     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGEGVYAFSELQAQLERGALARLE     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT     VWSAHGRAFT	i i	í	1	l	Ì		DRVETTRRYRLSFAHPPSAEVEAIALATLHDR
LALEKANOELGIALDSWDLDFYTKRTQELHVD   NPSTVARPIDLAGSNSEHSRHWFKGQLHVDG    NPSTVARPIDLAGSNSEHSRHWFKGQLHVDG    OKLVHSI,FSSIMSTORSSNPNNVLKECNSKI   OKEGVPER,FPSIMSTORSSNPNNVLKECNSKI   OKEGVPER,FPSIMSTORSSNPNNVLKECNSKI   OKEGVPER,FPSIMSTORSSNPNNVLKECNSKI   OKEGVPER,FPSIMSTORSSNPNNVLKECNSTORKEGPV   AETHNFFTGVCPFSGATTGTGGRIRDVCCTG   RGAHVVAGTAGYCFGNLHRGYNGLVBCDG   GRAFNELGIQLPDGQREWIKFIMFSGGGS   MEADHISKEAPPEMBEVVKVGPYNGGGG   GAASSVQVQDDNTSILDPGAVQRGDPEMBC   KMRRVIRACVEAPKGNPICSLHDQGAGGNG   NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW   GAEVQESNALLIKESPNDFLTHJDQGAGGNG   NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW   GAEVQESNALLIKESPNDFLTHJDQGAGGNG   NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW   GAEVQESNALLIKESPNDFLTHJDQGAGGNG   NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW   GAEVQESNALLIKESPNDFLTHJDQGAGGNG   NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW   GAEVGESNALERSVIGAUERCHRIPGAVASKRY   DAYLOKARING STORT			l				
NPSTVEATDLAQSNSEHSRHWFFKQQLHVDG (ALMHSLFESIKSTQESSNPMLYECDNSSA (ALMSLEFESIKSTQESSNPMLYECDNSSA) (ALMSLEFESIKSTQESSNPMLYECDNSSA) (ACKEVERLRFEDFTERSRFQQQQGLRHVYF) AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGYCFGNLHIRGYFNLFWEDLS) (APPONFARRLEVAIGASNGASDYGNKFGESIKSTQ CYPONFARRLEVAIGASNGASDYGNKFGESIKSTQ CYPONFARRLEVAIGASNGASDYGNKFGESIKSTQ MEADHISKEAPEPOMEVVKVGGPVTRICOG GAASSVQVQQDNTSOLDFGAVQRODPEMEC KMMRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAITYSRFQLGDFINALEIW GAEVQESNALLIRSPNRDFLTHVSARRCFA CFVGTTTGDRRIVLVDDRECYPTLAALEIW GAEVQESNALLIRSPNRDFLTHVSARRCFA CFVGTTTGDRRIVLVDDRECYPTLAALEIW GAEVQESNALLIRSPNRDFLTHVSARRCFA CFVGTTTGDRRIVLVDDRECYPTLAALEW GAEVQESNALLIRSPNRDFLTHVSARRCFA MQPLALPPELSVQAQQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSILDPKVAA RLAVBAALTILVFALVTDLREVSGNWM WAAKLPGEGAALADACEANVAMAALGV VDGGKDSLSMAARVGTETVRASCSNWM WAAKLPGEGAALADACEANVAMAALGV VCPDITATVTPDLRHPEGRGHLLVYVALSFGQ HRLGGTALAQCFSQQ,GEHPPPIDLPENLVXA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEW AFAGNCGGQVDVPVRVDVSLYFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEGHTGE AGPHAMWRVSWAARVALTEEVFACHE AFAGNCGGGVDVPVRVDVSLYFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEGHTGE AGPHAMWRVSWAAVLTEGESGNGD EMADAFHLAGFEVWDVTMQDLCSGAGGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFF RAGGAELRRFRKRPDTTSLGVCNGCQLLAHLG WVGGDPNEDAAEMGPDSQPARRGLLLHNL SGRYESRWASVRQFGPALMIRGMEGGAVLI VWSAHGEGYVAFSSPELQAQIBAGGLAFILL SGRYESRWASVRQFGPALMIRGMEGGAVLI VWSAHGEGYVAFSSPELQAQIBAGGLAFILLSHDQVADRYQWSYTLTT PWLOLFNARNVTLEGSC  GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKXLIYVHLVIVLIT SGRWESRWASVRQFGPALMIRGMEGGAVLI VGSMEMPFAAQQLYRRIGKAYRLIFGFSAVS MYLLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1345 2695 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAGGAVLI VGSMEMPFAAQQLYRRIGKAYRLIFGFSAVS MYLLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK TQYSWILLDSFRTTIPQLAEALARQUDTVIN NVPFEVIKQRLTARWHPASGRVYNTEFNPPK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYQDTVIN NVPFEVIKQRLTARWHPASGRVYNTEFNPPK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYNTEFPPK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYNTEFPPK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYNTEFPPK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYNTEFPFK TVGIDDLTGGEPLQREDDKPFTVIRRLGSTYNTEFPK TVGIDDLTGGEPLQREDDKPFTVIRRL		ļ	1	1			
OKLVHSLFSIMSTQESSNPNVLKFCONST   OKGVPRIRPEDPTRPSRPQQQGLRHVYF    AETHINPPTGVCPFSGATTGTGGRIRDVCGG    AETHINPPTGVCPFSGATTGTGGRIRDVCGG    AETHINPPTGVCPFSGATTGTGGRIRDVCGG    GRAHVVAGTAGVCFGHLHIPGYSLPWEDIS    OYPONFARPLEVAIEASINGASDYGNKFGEPV    LAGPARSLGQLPDGQRREWIKPIMPSGGIGS    MEADHISKEAPEPGMEVVKVGGVTVRIVGVG    GAASSVQVQDDITSIDLDFGAVQRGDFBME    KINNEVIRACVEAPKGNPICSLHDQGAGGING    NVLKELSDPAGAIIYTSRFQLGDFTLNALEIW    GAEVQESNALLIKESPNDPLTIADQGAGGING    NVLKELSDPAGAIIYTSRFQLGDFTLNALEIW    GAEVQESNALLIKESPNDPLTIADQGAGGING    NVLKELSDPAGAIIYTSRFQLGDFTLNALEIW    GAEVQESNALLIKESPNDPLTIADQGAGGING    NVLKELSDPAGAIIYTSRFQLGDFTLNALEIW    GAEVQESNALLIKESPNDPLTIADQGAGGING    NVLKELSDPAGAIIYTSRFQLGDFTLNALEIW    GAEVQESNALLIKESPNDPLTIADQGAGGING    PTPPTPYDLELEWVLGKMPRKEFFLQRKPP    MLQPLALPPGLSVRQALGERVLRIPAVASKRY    LTNKVDRSVGGLVAQQQCVCPIPLADVA    VVALSHEELIGAATALGEDPVKSILDPKVAA    RLAVABALITILVALVTDLADVCSGNWM    WAAKLPGGAALADACEAWAVAMAALGV    VVALSHEELIGAATALGEDPVKSILDPKVAA    RLAVABALITILVALVTDLAVALVTDLAVASKRY    WAAKLPGGAALADACEAWAVAMAALGV    VVGGKDSLSMAAVGTETVRASCSIVISAV    VCPDITATVTPDLKHPPGGAHLLYVALSPQQ    HRLGGTALAQCTSQLGEPPDLDLPENLYSCOGULTALEM    AFAGNICGLQVDVFVRRVDVISVLFAEEBGSLVISAV    VCPDITATVTPDLKHPPGGGLPTDLDLPENLYSCOGULTALEM    AFAGNICGLQVDVFVRRVDVISVLFAEEBGSLVISAV    VCPDITATVTPDLKHPPGGGLPTDLDLPENLYSCOGULTALEM    AFAGNICGLQVDVFVRRVDVISVLFAEEBRGSVISAVAAVTFIFFH    AGGALRAFRKRVSVNGAVVLEEPVGELRALWEE    TSPQLDRQLGREPQVADLCSGAIGLDI FRQVAFVGGFSYADVLGSAKGWAAAVTFIFH    RAGAERRFKRKPDTFSLGVCNGCQLLALIG   WVGGDPNBDAAEMGPDSQAARGGALIRHINNI    SGRYESRWASVRYGPQPALIRHEGGAVLE    WVSAHGGGYVAFSSPELQAQIPAGAICSCDDR    HLAVMPHPERAVRYWQWAWRPPPFDTLTTS    PWLQLFINARNMTLEGGSC    VRGPWAGTMARSRGGRRAAGTDGSDFQHRI    RVAMHYQMSVTLKYEERKLIYVYHLLVWLIVILIVILIVILIVILIVILIVILIVILIVILIVI	i i	]		ļ			
IQGKEVRFLRPEDPTRSKRQQQGERHVVFI ARTHNFPTGVCPFSGALTIOTGGRIRDVCCTG RGAHVVAGTAGVCFGNLHIRGYNLPWEDLS QVPGNFARPLEVAIRASNGASDYGNKFGEPV LAGFARSLGI.QLPDGQREEWIKPIMFSGGIGS MEADHISKEAPEPGMEVVKVGGPVYRIGVG GAASSVQVQGDNTSDLDFGAVQRGDFEMEC KMMRVIRACVEAPKGNFICSLHDQGAGGNG NVLKELSDPAGAUTYSRFQLGDFTLNALEIW GAEYQESNALLLRSPNRDFLTHVSARRCPA CFVGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPTTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERCPVRRNGQGDAP PTPPTTYDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERCPVRRNGQGDAP PTPPTTYDLELEWVLGKMPKKEFFLQRKPP MLQPLALPPGLSVHQALERGANAGAGAACAWAVMAALGV, VVALSHEELIGAATALGEPCVRRNGQGDAP RHAPASSVGGLVAQQCCVGPLQTPLADVA VVALSHEELIGAATALGEPGVSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALAACAMVAVMAALGV, VDGGCDSLSMAARVGTETVRAPGSLVISKY, VCPDITATVTPDLKHFBGRGHLLYVALSFQQ HRLGGTALAQCFSQLCEHPDLDLPENLVRA FSITQGILKDRLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVVDVSVLFAEEFGL LEVQEPDLAQVLKRYRDAGHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEFRCVAEEGRGLRRMGPSVC LEVGEPDLAQVLKRYRDAGHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEFRCVAEEGRGLRRMGPSVC JLPTTPFKASVPREPGGPSPVALREGGSNGDID EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAAVTFH RAGABLRRFRKRPDTTSLCVMGCQLLAHLG WVGGDPNEDAAEMGPDSQAAGCSCDGG HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARSWTLGSGS ADDGGNFGVSKSTLYHLVTWLLIJA AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFRNNISYLVLSNISMGIFSJAFLI VGSSMEMFFAAQQILYHRGKAYRFLFGRSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK AKMSVGHLRLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFRRNNISYLVLSNISMGIFSJAFLI VGSSMEMFFAAQQILYHRGKAYRFLFGRSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK DDFVRCGPGSAAMGASARLLRAVINGAPGS GGVLSKAFH DDFVRCGPGSAAMGASARLLRAVINGAPGS GGVLSKAFH TVGIDDLTGEPLIQREDDKEFTVIRKLAVED QTKPVLEYYQKKGVLLETFSGTETNKUMPYVS AFLQTKVPQKKGVLETFSGTETNKUMPYVS AFLQTKVPQKKGVLETFSGTETNKUMPYVS AFLQTKVPQKKGVLETFSGTETNKUMPYVS AFLQTKVPQKSQKASVTP ************************************			1				
AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGVCFGNLEFVILDVEDLS QYPGNFARPLEVAIRASNGASDYGNKFGEPV LAGFARSLGI,QIPDGQREWIKPIMFSGGGG MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEC KNMRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAITYTSRFQLGDFTLANLEIW GAEYQESNALLLRSPNRDFILTHVSARBCPA CFVGTITGDRRIVLVDDRECPVRRNGGDAP PTPPPTPVDLELEWVLGKMPEFFLQRKPP M.QPLALPPGLSWHQALERVIRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALDAACEAMVAWMAALGV, VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKIPEGGGALTVALLEM HLGGTALAQCFSOLGEHPPDLDLPENLVRA FSITGGLLKDRILLSGHDVSGLVTCLLEM FSTQGLKDRILLSGHDVSGLVTCLLEM FSTGGLKDRILLSGHDVSGLVTCLLEM FSTGGLKDRILLSGHDVSGLVTCLLEGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEFRCVAEEERGLRRRMGFSVC LPPTPFKASVPREPGGPSRYALLREEGSNGDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPPSVALREEGSNGDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGSNGCDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPGSAGGLALAL SGRYSSRWASVRVOPGPALMLRGMEGAVLB VWSAHGGGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSFGGVAGGCSCDGE HLAVMPHERAVRP WQWWRPPPDTILTTS PWLQLFINARNWTLEGSS  GRYSSRWASVRVOPGPALMLRGMEGAVLB VWSAHGGGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSFGGVAGGCSCDGE HLAVMPHERAVRP WQWWRPPPDTILTTS PWLQLFINARNWTLEGSS WGSWSTLKYBEKKLIYVHLVTWLLL AKMSVGHLRLLSHDQVAMPYQWEYPLLSI LPSLLGLLSFPRNNISTLVLSMSMGLFSLAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS GEGVLAKAFIDOGKLIPDDWMTRLALHELNN TQSSWLLDGFRTLPQAEALDRAYQIDTVNN NPFEVIKQRLTARWHPASGRVYNEENPPR TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GKTVLSETYTQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLE		i			}		QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGVCFGNLEFVILDVEDLS QYPGNFARPLEVAIRASNGASDYGNKFGEPV LAGFARSLGI,QIPDGQREWIKPIMFSGGGG MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEC KNMRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAITYTSRFQLGDFTLANLEIW GAEYQESNALLLRSPNRDFILTHVSARBCPA CFVGTITGDRRIVLVDDRECPVRRNGGDAP PTPPPTPVDLELEWVLGKMPEFFLQRKPP M.QPLALPPGLSWHQALERVIRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALDAACEAMVAWMAALGV, VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKIPEGGGALTVALLEM HLGGTALAQCFSOLGEHPPDLDLPENLVRA FSITGGLLKDRILLSGHDVSGLVTCLLEM FSTQGLKDRILLSGHDVSGLVTCLLEM FSTGGLKDRILLSGHDVSGLVTCLLEM FSTGGLKDRILLSGHDVSGLVTCLLEGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEFRCVAEEERGLRRRMGFSVC LPPTPFKASVPREPGGPSRYALLREEGSNGDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPPSVALREEGSNGDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGSNGCDI PRGVAFVGGFSYADVLGSAKGWAAAVTFEIP RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPGSAGGLALAL SGRYSSRWASVRVOPGPALMLRGMEGAVLB VWSAHGGGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSFGGVAGGCSCDGE HLAVMPHERAVRP WQWWRPPPDTILTTS PWLQLFINARNWTLEGSS  GRYSSRWASVRVOPGPALMLRGMEGAVLB VWSAHGGGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSFGGVAGGCSCDGE HLAVMPHERAVRP WQWWRPPPDTILTTS PWLQLFINARNWTLEGSS WGSWSTLKYBEKKLIYVHLVTWLLL AKMSVGHLRLLSHDQVAMPYQWEYPLLSI LPSLLGLLSFPRNNISTLVLSMSMGLFSLAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS GEGVLAKAFIDOGKLIPDDWMTRLALHELNN TQSSWLLDGFRTLPQAEALDRAYQIDTVNN NPFEVIKQRLTARWHPASGRVYNEENPPR TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GKTVLSETYTQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLETFSGTETINKIPPYY TVGIDDLTGEPLIQREDDKPETVIKRLKAYED GTKPVLEYYQKKGVLE		1	ł	1	ł	}	IOGKEVRFLRPEDPTRPSRF0000GLRHVVFT
RGAHYVAGTAGVCFGNLHIEGYNLPWEDLS QVPGNFARPLEVAIRASNASDYGNKFGEPY LAGFARSLGI.QLPDGQRREWIKPIMFSGGIGS MEADHISKEAFEFGMEVVR.VGGPVYTGIVG GAASSVQVQGDNTSDLDFGAVQRGDFEMEC KMMRVIRACVFAPKGNFICSLHDQGAGONG NVLKELSDPAGAUTYSRFQLGDFTI.NALEIW GAEYQGSNALLLRSPRIDLTHVSARRCCPA CEVGTITIGDRRIVI.VDDRECVRINGQGDAP PTPPTTYDLELEWVLGKMPRKEFFLQRKPP MLQFLALPPGLSVHQALERVLRLPAVASKRY LTINKVDRSVGGI.VAQQCCVGFLQTTPLADVA VVALSHEELIGAATALGEQFVKSLLDPKVAA RLAVAEALTINLVFALVTDLRDVKCSGNVWM WAAKLFGEGAALADACEAMVAVMAALGV. VDGGKDSI.SMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKHFEGRGHLLYVALSPQQ HRLGGTALAQCFSQLGGHPPDLDEPNLVRA FSITQGLIKGRLLCSGHEVSDGGLVTCLLEW AFAGNCGLQVDVPVPRVDVLSVLFAEBEGLX LEVQEPDLAQCHLRYRRDAGHCLELGHTGE AGPHAMVRSVNGAVVLEEPVGELRALWEI TSFQLDRLQAFPRCVAEEFGGLRERMGPSVC LPFTFFKASVPREPGGPSPRVAILREGSNOD EMADAFHLAGFEVWDVTMOLGSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTTSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSOPARFGLLENHNL SGRYESRWASVRVPGFBALMLREGGAVLI VWSAHGGGVAFSSPELQAQIEARGLAPLHY ADDDGMPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHERA NRFWQWAWRPPFFDTLTTS UVWSAHGGGVAFSSPELQAQIEARGLAPLHY ADDGMPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHERA NRFWQWAWRPPFFDTLTTS UVWSAHGGGVAFSSPELQAQIEARGLAPLHY ADDGMPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHERA NRFWQWAWRPPFFDTLTTS UVSAHGEGVAFTLKYHLIVITULLI AKMSVGHLRLISHDQVAMPYQWEYYPLISI LPSILGILSFPRNNSYLVLSMSMGFSIAPL YGSMEMFPAAQQLYRHGKAYFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS GEGVLAKAFIDQGKLIPDDVMTRLALHELNT TQSSWLDDGFFRTLPQAEALDRAYQDDTVNN NPFEVUKQRLTARWHPASGRVYNERNPFK TVGIDDLTGEPLIQREDDKPTTVKRLKAYED GTKPVLEYYQKKGVLEITSGTETNKUPYYV NNPFEVUKQRLTARWHPASGRVYNERNPFK TVGIDDLTGEPLIQREDDKPTTVKRLKAYED GTKPVLEYYQKKGVLEITSGTETNKUPYYV AFLGITKVPQRSQLASVTP			1				
QYPGNFARPLEVAIRASNGASDYGNKFGEPV LAGPARSLGJ.OJPDGQRESHPMFMGGGGGS MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVGRGDPEMEC KMMRVIRACVEAPKGNPICSLIDDGAGGNG NVLKELSDPAGAITYTSRFQLGDFTLANLEIW GAEVQESNALLLESPNDRI-THVSARBCCPA CFVGTITGDRRIVL-VDDRECPVRRNGGDAP PTPPTPYDLELEWVLGKMPKEFFT,ORKPP ML.QPLALPPGLSWNGALBRVRLEPAVASKRY LTNKVDRSVGGLVAQQQCPGJOTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNVM WAAKLPGEGAALDAACAGAMVAVMAALGV, VOPGTATVTPDLKIPFGGRGHLLVVALSFQQ HRLGGTALAQCFSOLGEHPPDLDLPENLVRA FSITTGGLLKDRILCSGHDVSGLVTCLLEGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEPKCVAEEERGLRRMGPSVC LPPTPFKASVPREFQGFSRVALLREEGSNGDJ EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFFIF RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGGSNGD EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFFIF RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGGSNGD EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFFIF RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGGSNGD EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFFIF RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGGSNGD EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFFIF RAGAELRFRKRPDTFSLGVCNGCQLALLG WVGGDPNEDAAMGPSVALREEGSSNGD EMADAFHLAGFEVWDVTMQDLCSGAGLDT FRGVAFVGGFSYADVLGSKGWAGTSCADGE HLAVMPHERAVRPWQWWWPPPFDTLTTS FWLQLFINAMWTLGSGS GGTVSSRIVANVTLKSGEKLHYVHLVTWLLLL AKMSVGHLRLLSHDQVAMPYQWEYPLLSI LPSLLGLLSFPRNNIS'LVLSMSMGLFSLADL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS GEGVLAKAFIDOGKLIPDDWMTRLALHELN TQSWLLDGFFRTLPQAEALDRAYQIDTVIN NVFFEVUKQRLTARWHPASGRVYNEENPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYEL GTKPVLEYYQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKKGVLETFSGTETINGWPTYV AFLQTKVPQKGVLETFSGTETINGWPTYV AFLQTKVPQKGVLEDNTMGE	1	!	1	1		ļ	
LAGFARSLGLQLPDGQREWKPMPRSGGIGS MEADHISKAPPPGMEVKGGPVYNGVG GAASSVQVQGDNTSDLDFGAVQRGDPEMEC KMWRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRQLGDPTLNALEIW GAEYQESNALLIRSPNRDFLTHVSARRCCA CEVGTTIORDRIVLVDDRECPVRRNGQGDAP PTPPTPVDLELEWVLGKMPKKEFFLQRKPP MLOPLALPPGLSVPHOALERVLRIPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQEVKSLIPKVVAA RLAVAEALTNLFALVTDLRDVKCSGNWM WAAKLPGEGAALDACEAMVAVMAALGV VOBGKDSLSMARAVETVRAPGSLVISAY/ VCPDITATVTPLKHPPGRGHLLYVALSPGQ HRIGGTALAQCPSQLGETVRAPGSLVISAY/ VCPDITATVTPLKHPPGRGHLLYVALSPGQ HRIGGTALAQCPSQLGETVRAPGSLVISAY/ VCPDITATVTPLKHPPGRGHLLYVALSPGQ HRIGGTALAQCPSQLGELRALWEE AFAGNCGLQVDVPVPRVDVLSVISAY/ VCPDITATVTPLKHPPGRGHLLYVALSPGQ HRIGGTALAQCPSQLGELRALWEE TSPQLDRLQAVLKRYKDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSPQLDRLQAFRCVAEEERGLRERMGPSVC LPPTTPKASVPREPGGPSRVALIREEGSNGD) EMADAPHLAGFEWWDVAGDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDIPHEDAAEMGPDSQPARPGLLLRHNL SGRYESRWASVRVGPOPALMLRGMEGAVLE VWSAHGEGVYAFSSPLGVCNGCQLLALLG WVGGDIPHEDAAEMGPDSQPARPGLLLRHNL SGRYESRWASVRVGPOPALMLRGMEGAVLE VWSAHGEGVYAFSSPLGVCNGCQLLALLG WVGGDIPHERAVRPWIPEGSPGGVAGGCSCDGG HLAVMPHPERAVRPWWWARPPPFDTLTTS PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYLSIKKLINYVHLVIWLLIV AKMSVGFILRLISHDQVAMPYQWPYPTLSI LPSLLGLLSFPRNINSTVLNSMISMGFSAPLI VGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFIS QEKKHK TQYSWLLDGGFRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIPASGRVYNENPYRK TVGIDDLTOPELIQREDDKPETVKRLKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYED QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKKGVLETFSGTETNKRKAYPE QTKPVLEYYQKGVLETFSGTETNKGLENTNGES	İ						
MEADHISKEAPEPGÜVYKIGVGG GAASSVQQODNTSDLDFGAVQROPPEMEC KMNRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRRQLGDFULNALEIW GAEYQESNALLIRSRNDFLTHVSARERCPA CFVGTITGDRRIVLVDDRECPVRRNQGDAP PTPPTPTDLELBWLQRMRKEFFLGRKPP MLQPLALPPGLSVHQALERVLRIPAVASKR LTNKVDRSVGGLVAQQQCVGPLQTPLADVA WAAKLPGGAALADACEAWAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALADACEAWAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALADACEAWAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALADACEAWAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALLAVAGNAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALLAVAGNAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKPBGGALLAVE AFAGNGCLQVDVPVPRVDVLSVLASERGG HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGALVTCLEEM AFAGNGCLQVDVPVPRVDVLSVLASERGG LEVQEPDLAQVLKRYRDAGHLCLEGHTGG AGPHAMVXSVNGGAVULEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGERALWE TSFQLDRLQAVERVDVALEPVGGLLALGU WVGGDPNEDAAEMGPDSQPARFGLLRHNL SGRYESRWASVRYPGGPSLLRHNL SGRYESRWASVRYPGGPSLLRHNL SGRYESRWASVRYPGFALLRHEN SGRYESRWASVRYPGFALLRHNL SGRYESRWASVRYPGFALLRHNL SGRYESRWASVRYPGFALLRHEN SGRYESRWASVRYPGFALLRHNL SGRYESRWASVRYPGGALLANGGESCDGE HLAVMPHPERAVERVWAWRPPPPDTLTTS PWLQLFINARNWTLEGSC GVRGFWAGATMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLTYVHLVWLLLA AKMSYGHLRLLSHDQVAMPYQWEYPYLLSI LFSLLGLLSFPRNNISYLVLSMMSGLFSLAPLL YGSMEMFPAAQQLYRHGKAYRFLFGFSAYS MYLVLVLAQVHAWQLYYTLKYEIKKLTYVHLTWLLLA AKMSYGHLRLLSHDQVAMPYQWEYPYLLSI LFSLLGLLSFPRNNISYLVLSMMSMGLFSLAPLL YGSMEMFPAAQQLYRHGKAYRFLFGFSAYS MYLVLVLAQVHAWQLYYTLASH TYGSMCLTGAGAGAARLTRAVIMGAPGS GKGTVSSRTTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTTALAHELKN TQYSWLLDGFPRTLPQARALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRYVINENNPYS TVGIDDLTGFPLTQREDDKPTVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYY AFLQTKVPQRSQKASVTP KHRQEDTAALDMAPELIBHTGPMCLIENTNGEI			1	1			l •
GAASSVQVQGDTSDLDFGAVORGDPEMEC KMNRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRQLGDPTLNALEIW GAEYQESNALLIRSPNRDPI.THVSARRCCPA CEVGTITIORRIVIV.VDDFCVPRRNQGDAP PTPPPTPVDLELEWVLGKMRKEEFILORKPP MLOPLALPPGLSVHQALERVIRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQEVKSLDPKVAA RILAVABALTNLVFALVTDLRDVKCSGNVM WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMARAVETVRAPGSLUDPKLVAA FSITQGLLKSSLMARAVETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDIPENLVRA FSITQGLLKDRLCSGHDVSDGGLVTCLLEM APAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEFRCVAEEERGLRERMGPSYC LPPTPFKASVPREPGGPSPVAULREEGSNGD EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSVADVLGSGAIGLDT FRGVAFVGGFSVADVLGSGAIGLDT FRGVAFVGGFSVADVLGSGAIGLDT FRGVAFVGGFSVADVLGSGAIGLADL WVGGDPNEDAAEMGPDSQPARPGCLLRHNU SGRYESRWASVRVPGGPALLIRHGKAAVLTV VWSAHGEGYVAFSSPELQAQIEAGLAPLHV ADDDGNFTEQYPLNPNGSFGGVAGICSCDGG HLAVMPHPERAVRPWQWWWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYYHLVIWLLL LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGLLSFPRNNISYL VLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWWYQWEYPYLLSI LPSLLGGFRLTLRGADARAYQIDTVIN NVPFEVIKQRLTARWHPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYQIDTVIN NVPFEVIKQRLTARWHPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYQIDTVIN NVPFEVIKQRLTARWHPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYQIDTVIN NVPFEVIKQRLTARWHFPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYGIDTVIN NVPFEVIKQRLTARWHFPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYGIDTVIN NVPFEVIKQRLTARWHFPASGRYVNEENPIPK TVGIDDLTGFFLLTRGADARAYGIDTVIN NVFEVIKGRLTARWHFPASGR		İ	l	1			
KMMRVIRACVEAPKGMPICSLINGGAGGING NVLKELSDPAGAINTSRPQLGDPTLNALEW GAEYQESNALLIRSPNRDICTHYSARERCPA CEVGTITGDRRIVLYDDRECPVRRNQGDAP PTPPTTEVDLELEWVLGKMPRKEFFLQRKPP MLQPLALERVLRLPAVASKRY LTINKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVABALTNILVFALVTDLRDVKCSGNVM WAKAKLPGGAALADACEMVAVMAALGV. VDGGKDSLSMAARVGTETVRAPOSLVISAY, VCPDITATYTPLKHPGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDILPENLVRA FSITQGLLKDRLLCSGHDVSDGLVTCLLEM AFAGNCGLQVDVPVRVDVLSVLFAEBPGLL LEVQEPDLAQVLKRYRDAGHLLTVLLEGHTGE AGPHAMVAVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEPRCVAEEERGLRERMGPSVC LPPTTPKASVPREPGGFSPRVALIREGGSNGDI EMADAFILAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAVTTPH RAGAGELRRFKRPDTSTSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNU SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQICARGLAPLHY ADDDGNPTEQYPLNPNGSPGVAGICSCDGE HLAVMPHPERAVRPWQWAWRPPPPDTLTTS PWLQEFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLTYYHLVWLLLA AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSILGLISFPRNNISYLVLSMISMGLFSLAPLL YGSMEMFPAAQQLYRHGKAYRFLGFSAVS MYLVLVLAQVHAWQLYYKKLLDSWFTS QEKKHK TOPST THE STANDARV GEGVLAKAFIDQGKLIPDDVMTRALHELKNI TYSPWLLQHARAFIDGGKLIPDDVMTRALHELKNI TYSPWLLDGFPRTFGFSAVS MYLVLVLAVQVHAWQLYRKKRIALBELKNI TYSPWLLDGFPRTFGFSAVS MYLVLVLAVQVHAWQLSVRIEFFGFSAVS MYLVLVLAVQVHAWQLSVRIEFFGSAVS MYLVLVLAVQVHAWQLSVRIEFFGSAVS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRALHELKNI TYSPWLLDGFPRTFLQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS QEKKHK TYSPWLAVGAGALARAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABA		]	)	]	}	į	MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
KMMRVIRACVEAPKGMPICSLINGGAGGING NVLKELSDPAGAINTSRPQLGDPTLNALEW GAEYQESNALLIRSPNRDICTHYSARERCPA CEVGTITGDRRIVLYDDRECPVRRNQGDAP PTPPTTEVDLELEWVLGKMPRKEFFLQRKPP MLQPLALERVLRLPAVASKRY LTINKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVABALTNILVFALVTDLRDVKCSGNVM WAKAKLPGGAALADACEMVAVMAALGV. VDGGKDSLSMAARVGTETVRAPOSLVISAY, VCPDITATYTPLKHPGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDILPENLVRA FSITQGLLKDRLLCSGHDVSDGLVTCLLEM AFAGNCGLQVDVPVRVDVLSVLFAEBPGLL LEVQEPDLAQVLKRYRDAGHLLTVLLEGHTGE AGPHAMVAVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEPRCVAEEERGLRERMGPSVC LPPTTPKASVPREPGGFSPRVALIREGGSNGDI EMADAFILAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAVTTPH RAGAGELRRFKRPDTSTSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNU SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQICARGLAPLHY ADDDGNPTEQYPLNPNGSPGVAGICSCDGE HLAVMPHPERAVRPWQWAWRPPPPDTLTTS PWLQEFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLTYYHLVWLLLA AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSILGLISFPRNNISYLVLSMISMGLFSLAPLL YGSMEMFPAAQQLYRHGKAYRFLGFSAVS MYLVLVLAQVHAWQLYYKKLLDSWFTS QEKKHK TOPST THE STANDARV GEGVLAKAFIDQGKLIPDDVMTRALHELKNI TYSPWLLQHARAFIDGGKLIPDDVMTRALHELKNI TYSPWLLDGFPRTFGFSAVS MYLVLVLAVQVHAWQLYRKKRIALBELKNI TYSPWLLDGFPRTFGFSAVS MYLVLVLAVQVHAWQLSVRIEFFGFSAVS MYLVLVLAVQVHAWQLSVRIEFFGSAVS MYLVLVLAVQVHAWQLSVRIEFFGSAVS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRALHELKNI TYSPWLLDGFPRTFLQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS QEKKHK TYSPWLAVGAGALARAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABALDRAYQUTVIN NVPFEVIKQRLTARWIHPASGRVYNEFNPYS TYGDDLTGFPLTQABA				1			GAASSVOVOGDNTSDLDFGAVORGDPEMEO
NVLKELSDPAGAIIVTSRFQLGDPTLNALEIW GAEYQESNALLIRSPNRDFLTHYSARRCPA CFYGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPTPTPUDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVPLQALERVLRIPAVASKRY LTNKVDRSVGGLVAQQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV/ VDGKGSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSOLGEHPPDLDIPENLVRA FSITQGLIKGRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVJSVLFAEEPGIL LEVQEPDLAQVLRKYRDAGIHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSPCLDRLQAEPRCVBEERGIRERMGPSVC LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRGVAPVGGFSYADVLGSAKGWAAAVTFHF RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGPDREDAAEMQPDSQPAPPGLLIRHNI SGRYESRWASVRVOPGPALMLRGMEGAVLE VWSAHGEGYVAFSSPELQAQICARGLAPLHV ADDGDNPTEQVPLNPNGSPGGVAGICSCDGE HLAVMPHERAVRPWQWARPPPDTTLTTS PWLOLFINARNWTLEGSS GYGGRWAGTMASRAGFRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIVYHLVIVLUTL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLGLISFPRNNISVLVLSMISMGLFSLAPLL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVAAVVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPGSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDVMTLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRYNIERNPIPK VGGDDLTGFPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENGTETNKIWPYV3 AFLQTKVPQRSQKASVELTENTNGEI	1		1	1		1	
GAEYQESNALLIRSPNRDĒLTHVSARERCPA CFVGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPTPYDLELEWVIGKMPRKEFTLQRRPP MLQPLALPPGLSVHQALERVLRIPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTILRADVKCSGNWA WAAKLJPEGAALADACEAMVAVMAALGV. VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKBEPGGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEBFGLL EVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEBERGLRERMGPSYC LPPTPFKASVPREPGGPSPRVAULREEGSNGDD EMADAPHLAGFEV WDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHE RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVROFGPALMLRGMEGAVLF VØSAHGGGVVAFSSPELQAQHEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHERAVRPWQWAWRPPPDTLTTS PVQLEFINARNWTLGSC GVKGPWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVWLLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFFRNNISYLVLSMISMGCFSIAPLL YGSMEMFPAAQQLYRHGKAYFFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1346 2696 A 10398 1 718 DDFVRCGFQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRYYNEENDPK TVGDDLTGEPLIGREDDKPETVIKRIKAYED QYKRPULEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP TVGDDLTGEPLIGREDDKPETVIKRIKKAYED QYKRPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTVP	[	[	1	1		,	
CFYGTITGDRRIVL/DDRECVPRRNGQGDAP PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRIP,AVASKRY LTNKVDRSVGGLVAQQCOVGPLQTTLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNVM WAAKLPGEGAALADACEAMVAVMAALGV, VDGGKDSLSMARNCTETVRAPGSLVISAY, VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITGGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLN LEVQEPDLAQVLKRYRDAGHCLELGHTIGE AGPHAMVRVSVNGAVVLEEPYGELRALWEI TSGLDDRLQAEPRCVAEERGLRERMGPSVC LPPTPPKASVPREPGGPSPRVALLREEGSNGDI EMADAFILAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPREDAAENGDSQPARFGLLRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLH VWSAHGEGYVAFSDEQLAGHLAGHLAY ADDGONPTEQVPLNPNGSPGGVAGICSCDGR HLAVMPHERAVRPWQWAWRPPPFDTLTTS PVQLEINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVANHYQMSVTLKYEIKKLIYVHLVIVLIX AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFFRNNISVLVSMSMGLFSLAPLL YGSMEMFPAAQQLYRHGKAYFELFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGFQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRILPQAEALDRAYQIDTVIN NVPFEVKQRLTARWIHPASGRYYNIENPIPK VGMDDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTIKYPQRSQKASVTYI TVGMDDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTIKYPQRSQKASVTYI TVGMDDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTIKYPQRSQKASVTYI TVGMDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGREDDKPETVIKRIKAYED QTKPVLEYYQKRGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGREDBKPETVIKRIKAYED QTKPVLEYYQKRGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTYI TVGMDLTGEPLIGRETNKIMPT TVGEN			İ				
PTPPPTPVDLELEWVLGKMPRKEFFLÖRKPP MLQPLALPPGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEGPVKSLLDPKVAA RLAVAEALTNILVATGEGPVKSLLDPKVAA RLAVAEALTNILVATGEGPVKSLLDPKVAA RLAVAEALTNILVATGEGPVKSLLDPKVAA RLAVAEALTNILVATGEGPVKSLLDPKVAA RLAVAEALTNILVAGEGPVKSLLDPKVAA RLAVAEALTNILVAGEGPVKSLLDPKVAA RLAVAEALTNILVAGEGPVKSLLDPKVAA RLAVAEALTNILVAGEGPALADACEAMVAWAALGV/ VDGGKDSLSMAARVGTETVAPGSLVISAY/ VCPDITATVTPDLKHPEGGRIFLLYVALSFGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKORLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGIL LEVQEPDLAQVLKRYRDAGHCLELGHTGE AGPHAMYNSVNGAVVLEEPVGEIRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSVC LPPTFPKASVPREPGGPSPRVALREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSVADVLGSAKGWAAAVTFHF RAGAELRRFKKRPTSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESRWASVRVGPGPALMLRGMEGAVLE VWSAAHGEGYVAFSSPELQAGIEARGLAPLHY ADDGMPTEQYPLNPNGSSPGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTILTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDOSDFQHRI RVAMHYQMSVTLKSEKLITYVHLVTWLLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISTLVLSMISMGLFSIAPLL YGSMEMPPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1346 2696 A 10398 1 718 DDFVRCGFQSAAMGASARILRAVIMGAPGS GKGTVSSRTTHFELKHLSGDLLRDNMLRG EIGHLAKAFIDQGKLIPDDVMTRLAHLEKKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHFBAGRYYNIEFNPK TQGIDDLTGFEIQREDDKPSTVKRLAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVA AFLQTKVPQRSQKASVTF			l	i	İ	}	
MI.QPI.ALPPGLSVHQALERVLRI.PAYASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNI.VFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGW VDGGKDSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTPDLKHPEGGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLN LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMWRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREFGGPSRVALLREEGSNGD) EMADAFHLAGFEWWDVTMQDLCSGAIGLDT FRGWAPVGGFSVAULGSAGGWAAVTFHFT RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEADALGGPSORWAAAVTFHFT RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEADAAEMGPDSOPARRGLLRINIL SGRVESRWASVRVGPGPALMLREGSSCOCE HLAVWPHPERAVPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLLYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGYLAKAFIDQGLIPDDVMTRLALHELKIN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVKKQRLTARWHFASGRVYNIEFNPPK TGGIDDLTGEFLIQREDDKPETVIKRLAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	l			CFVGTITGDRRIVLVDDRECPVRRNGQGDAP
LTNKVDRSVGLVAQQCVGPLQTFLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNILVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV, VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLKDRILCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTIGE AGPHAMVRVSVNGAVVLEEPYGELRALWEE TSFQLDRLQAEPRCVAEERGLRERMOPSVC LPPTFPKASVPREGPSPRVALIREGSSNGD) EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGABLRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARFGLLIRHNL SGRYESRWASVRVQPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDGMPTEQYPLNPNGSPGGVAGICSCDGF HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPPSLIGLLSFPRNNISYLVLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRTTTHFELKHLSSGDLLRDNMLRG' EIGYLAKAFIDQGLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWHFPASGRVYNIEFNPK TVGIGDDLTGFELIQREDDKPETVKRRLAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYV) AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAAPEHHNTGPMCLIENTNGEI			1				PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
LTNKVDRSVGLVAQQCVGPLQTFLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNILVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV, VDGGKDSLSMAARVGTETVRAPGSLVISAY, VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLKDRILCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTIGE AGPHAMVRVSVNGAVVLEEPYGELRALWEE TSFQLDRLQAEPRCVAEERGLRERMOPSVC LPPTFPKASVPREGPSPRVALIREGSSNGD) EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGABLRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARFGLLIRHNL SGRYESRWASVRVQPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDGMPTEQYPLNPNGSPGGVAGICSCDGF HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPPSLIGLLSFPRNNISYLVLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRTTTHFELKHLSSGDLLRDNMLRG' EIGYLAKAFIDQGLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWHFPASGRVYNIEFNPK TVGIGDDLTGFELIQREDDKPETVKRRLAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYV) AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAAPEHHNTGPMCLIENTNGEI	ľ		İ	1	i		
VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV/ VDGGKDSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTDKHPEGRGHLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVRVDVLSVLFAEEPGLN LEVQEPDLAQVLKRYRDAGHCLELGHTGE AGPHAMVRVSVNGAVULEPVEGLRALWEI TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTPPKASVPREPGGPSFRVALREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFFKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVRVOPGPALMLRGMEGAVLF VWSAHGEGYVAFSSFELQAQIEARGLAPLHY ADDGNPTEGVYLNFNGSPGGVAGIGSCDGR HLAVMPHPERAVRPWQWAWRPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLY AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSILGLLSFPRNNISVLVLSMISMGIFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWTTS' QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLIRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGULAKAFIDQGKLPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP	ł	l	ł		1	ł	
RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKIPGEGAALADACEAMVAVMAALGV/ VDGGKDSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTPDLKHPEGGGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITGGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSVC LPPTFPKASVPREPGGPSPRVAILREEGSNGDD EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHINL SGRYESRWSVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSSPGVAGICSCDGR HLAVMPHPERAVPRWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIVILLI AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLISFPRNNISVLVISMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLAHELKNI TQYSWLLDGFPRTLPQAEALDRAYQDITVIN NVPFEVIKQRLTARWHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAPEHMTGFMCLIENTNGEI		ļ	<b>{</b>				
WAAKLPGEGAALADACEAMVAYMAALGV/ VDGGKDSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTPÜLKIPEGRGHLLYVALSPGQ HRLGGTALAQCPSQLGEHPPDLDLPENLVRA FSITOGLKKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTPPKASVPREPGGPSPRVAILREGGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRCVAFVGGFSYADVLGSAKGWAAAVTFHP RAGAELRFRKRPDITSLGYCNGCQLLALLG WVGGDPNEDAAEMGPDSOPARPGLLRIHNI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNYTEVPLNPNGSSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPFDTLTTS PWLQLFINARNWILGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLA AKMSVGHLRLLSHIDQVAMPYQWEYPYLLS LPSELGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGBPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPPEIHMTGPMCLIENTNGEI			1	l	ļ		ſ
VDGGKDSLSMAARVGTETVRAPGSLVISAY/ VCPDITATVTPDLKHPEGRGHLLYVALSRGJ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTIGE AGPHAMVRVSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEFRCVAEEERGLRERMGPSYC LPPTPFKASVPREPGGPSPRVAILREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDI FRGVAFVGGFSYADVLGSAKGWAAAVTFHE RAGAELRRFRKRPDTFSLGVGNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRINL SGRYESRWASVRVYGGPALMLRGMEGAVLI VWSAHGEGYVAFSSPELQAQIEAGGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHERAVRPWQWAWRPPPTDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGFQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRITLQAEALDRAYQUDTVIN NVPFEVIKQRLTARWHPASGRYYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  LT347 2697 A 10402 153 1969 KHRQENNALDMAPEHHMTGPMCLIENTNGEL	i i	1	,				
VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDFENLYMA FSITYGLLKDRILCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRYSVNGAVVLEEPVGELRALWEI TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTTPKASVPREPGGPSPRVAILREGGSNGDD EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLOSAKGWAAAVTFHF RAGAELRFRKRPDTTSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESR WASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHY ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGFRAAGTDGSDFQHRI RVAMHYQWSYTLKYEIKKLIYYHLVIWLLLX AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYFFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRITLQAEALDRAYQIDTVNI NVPFEVIKQRLTARWHPASGRVYNIEFNIPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEHHMTGPMCLIENTNGEI			l	1		Ì	WAAKLPGEGAALADACEAMVAVMAALGVA
HRLGGTALAQCFSQLGEHPPDLDILPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNGGLQVDVPVPRVDVLSVLFAEEPGLI LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVVIGAVVLEEPVGELRALWEI TSFQLDRLQAEPRCVAEEFGGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFHILAGFEVWDVTMQDLCSGAIGLDI FRGVAFYGGFSYADVLGSAKGWAAVTFHE RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPBSQPARPGLLRINI SGRYESRWASVRVGPGPALMLRGMEGAVLE ADDDGNPTEQVPLNPNGSPGGVAGICSCDGE HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLN AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGIFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVKRQRLTARWIHPASGRYVNIEFNPRK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFILQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPDHIMTGPMCLIENTNGEI			ļ	1			VDGGKDSLSMAARVGTETVRAPGSLVISAYA
HRLGGTALAQCFSQLGEHPPDLDILPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNGGLQVDVPVPRVDVLSVLFAEEPGLI LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVVIGAVVLEEPVGELRALWEI TSFQLDRLQAEPRCVAEEFGGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFHILAGFEVWDVTMQDLCSGAIGLDI FRGVAFYGGFSYADVLGSAKGWAAVTFHE RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPBSQPARPGLLRINI SGRYESRWASVRVGPGPALMLRGMEGAVLE ADDDGNPTEQVPLNPNGSPGGVAGICSCDGE HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLN AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGIFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVKRQRLTARWIHPASGRYVNIEFNPRK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFILQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPDHIMTGPMCLIENTNGEI		1	Į.	ı			VCPDITATVTPDLKHPEGRGHLLYVALSPGO
FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREPGGPSPPVAILREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRINNL SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGMPTEQYPLNPNGSPGGVAGICSCDGE HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNISYLVLSMISMGLFSLAPLI YGSMEMPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			l				l
AFAĞNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEPVGELRAL WEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREPGGPSPRVALLREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAPVGGFSYADVLGSAKGWAAAVTFHF RAGAELRRFKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRINNI SGRYESRWASVRVGPGPALMLRGMEGAVLLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGF HLAVMPHERAVRPWQMWRPPPFDTLTTS PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK TGYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPFK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKYPQKSQKLSVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		•	1	}	j	j	, · · · · · · · · · · · · · · · · · · ·
LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEPVGELRALWEIF TSFQLDRLQAEPRCVAEERGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RGAELRFFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNL SGRYESRWASVRVOPGPALMLRGMEGAVLLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC GVGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVAVQVHAWQLYYSKKLLDSWFTS' QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPEFVKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		i	1				l
AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEFRCVAEEERGLRERMGPSYCO LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARFGLLLRINI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCOGE HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC T345 2695 A 10396 65 642 GVRGFWAGTMASRAGFRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLL LPSLLGLLSFPRNNISYLVLSMISMGLFSLAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAQVHAWQLYYSKKLLDSWFTS' QEKKHK DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		l					AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV
TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTTPKASVPREPGGPSPRVAILREGSNGDD EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLIRHNIL SGRYESRWASVRVGPGALMLRGMEGAVLI VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLX AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVIAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLEFTSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	1	1	l	i			LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFILAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLV AKKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVX AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEHMTGPMCLIENTNGEI	1	1	ł			}	AGPHAMVRVSVNGAVVLEEPVGELRALWEE
LPPTFPKASVPREPGGPSPRVAILREEGSNGDI EMADAFILAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHF RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLV AKKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVX AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEHMTGPMCLIENTNGEI					ŀ		TSFOLDRLOAEPRCVAEEERGLRERMGPSYC
EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHE RAGAELRRFRKPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLE VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSSPGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG- EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKRVVLEYYQKKGVLETFSGTETNKIWPYVY AFFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			}				
FRGVAFVGGFSYADVLGSAKGWAAAVTFHE RAGAELRRFKKPPTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLIRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGFRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		ſ	1	ĺ			l
RAGAELRRFRKPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNI SGRYESRWASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFRRNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVA AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI						l	1
WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESR WASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HILAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			l	1	1	}	
WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESR WASVRVGPGPALMLRGMEGAVLF VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HILAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI						'	
SGRYESRWASVRVGPGPALMLRGMEGAVLE VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHIU RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI				l	1		WVGGDPNEDAAEMGPDSQPARPGLLLRHNL
VWSAHGEGYVAFSSPELQAQIEARGLAPLHV ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HILAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLLVIWLLLY AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			l	1	1		
ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		ł	i	l		Ì	
HILAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	1	1	1	
PWLQLFINARNWTLEGSC  1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	}		ł	ł		}	
1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRI RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			ł	1	Ī	1	
RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		L	ŀ	L_			PWLQLFINARNWTLEGSC
RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	1345	2695	Α	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFOHRE
AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		}	l		1	} - ·-	
LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			l				
YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	]	]	J	]	]	ł	
MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	1			
MYLVLVLAVQVHAWQLYYSKKLLDSWFTS' QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG' EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	1	1	1	
QEKKHK  1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	(		1	1	{	[	MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	i	1	]	
GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	1346	2606	<del> </del>	10200	1	710	
EIGVLAKAFIDQGKLIPDDVMTRLALHELKNI TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	1240	2090	۱^	10390	1	1 ′10	
TQYSWLLDGFPRTLPQAEALDRAYQIDTVINI NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	[		ľ	1	Ī	1	
NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			ł		<b> </b>	<b>!</b>	
NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI	}		1	l	ł	}	TOYSWLLDGFPRTLPQAEALDRAYOIDTVINL
TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			1	l			
QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP  1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI		-	·	ł . ·	- · · · · ·	- · -	
AFLQTKVPQRSQKASVTP   1347   2697 A   10402   153   1969   KHRQENNALDMAPEIHMTGPMCLIENTNGEI			l	1		1	
1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			ł	l	1		
1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEI			Į		l	1	AFLQTKVPQRSQKASVTP
	1347	2697	A	10402	153	1969	
I I VANPEALKILNAHOPVVVVAIVŒLYKTŒKSY	' '	'					VANPEALKILSAITQPVVVVAIVGLYRTGKSY
VALUE ENGLATED AND A CONTROL OF THE	L	<b></b>	Ь———	<u> </u>	L	<u></u>	THE PROPERTY OF THE OWNER OWNER OF THE OWNER OWNE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ľ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		<b>!</b>		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}		1	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł		peptide		/=possible nucleotide deletion, \=possible
<u> </u>				sequence		nucleotide insertion
						LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
}		ļ				PHPKKPEHTLVLLDTEGLGDVKKGDNQNDS
						WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY
		1				VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
		[				WTLRDFSLDLEADGQPLTPDEYLEYSLKLTQ
			•			GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
		ĺ				HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
1						FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
Ì			i			GDLPCMENAVLALAQIENSAAVQKAIAHYD
1	ļ	1				QQMGQKVQLPAETLQELLDLHRVSEREATEV
						YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
						QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
ł						PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
1		ĺ				QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
		•				VKAESAQASAKMVEEMQIKYQQMMEEKEKS
		j				YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
						QEQARVLKERCQGESTQLQNEIQKLQKTLKK
į.	1			1		TOTAL TOTAL TOTAL TOTAL
<u> </u>						KTKRYMSHKLKI
1348	2698	Ā	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
1348	2698	Ā	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV
1348	2698	A	10404	59	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
1349	2699	A	10409	59	1184	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM
						TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
1349	2699	A	10409	59	1184	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
1349	2699	A	10409	59	1184	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSDDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
1349	2699	A	10409	59	1184	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW

#### WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
  - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20

## (19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

**PCT** 

# (10) International Publication Number WO 01/57188 A3

(51) International Patent Classification<sup>7</sup>: C07K 5/00, A61K 39/395, C12Q 1/68

C07H 21/04.

(21) International Application Number: PCT/US01/03800

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) US 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US Filed on 09/496,914 (CIP) 3 February 2000 (03.02.2000)

US Filed on 09/560,875 (CIP) 27 April 2000 (27.04.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court. San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.: Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE. AG. AL. AM. AT. AU. AZ. BA. BB. BG, BR. BY. BZ. CA. CH. CN. CR. CU. CZ. DE, DK. DM. DZ. EE. ES. FI. GB. GD. GE. GH. GM. HR. HU, ID, IL, IN, IS, JP. KE. KG. KP. KR. KZ. LC. LK. LR. LS. LT. LU. LV. MA. MD. MG. MK. MN. MW, MX. MZ. NO, NZ. PL. PT. RO. RU, SD. SE. SG, SI, SK. SL. TJ. TM. TR. TT. TZ. UA. UG. US. UZ. VN. YU. ZA. ZW.
- (84) Designated States (regional): ARIPO patent (GH. GM. KE, LS, MW. MZ, SD, SL, SZ, TZ, UG, ZW). Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM). European patent (AT, BE, CH, CY, DE, DK, ES, Fl, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, Cl, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

- -- with international search report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
- (88) Date of publication of the international search report: 28 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

l/57188 A

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

### INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/03800

A. CLASSIFICATION OF SUBJECT MATTER									
IPC(7) :C07H 91/04; C07K 5/00; A61K 39/395; C12Q 1/68 US CL :536/23.1; 530/300; +2+/130.1; +35/6									
US CL :536/23.1; 530/300; +24/130.1; 435/6 According to International Patent Classification (IPC) or to both national classification and IPC									
B. FIELDS SEARCHED									
Minimum documentation searched (classification system followed	by classification symbols)								
U.S. : 536/23.1; 530/300; +2+/130.1; +35/6									
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched									
Electronic data base consulted during the international search (na STN, medline search terms: EST, expressed sequence tags	nme of data base and, where practicabl	e, search terms used)							
C. DOCUMENTS CONSIDERED TO BE RELEVANT									
Category* Citation of document, with indication, where app	propriate, of the relevant passages	Relevant to claim No.							
Map Based on Naturally Occurring I	VOLLRATH. D. et al. The Human Y Chromosome: A 43-Interval Map Based on Naturally Occurring Deletions. Science. October 1992. Vol 258. pages 52-59, see whole document.								
Further documents are listed in the continuation of Box C	See patent family annex.								
Special categories of cited documents:	"I" later document published after the inte								
"A" document defining the general state of the art which is not considered to be of particular relevance	the principle or theory underlying the	invention							
"E" carlies document published on or after the international filing date	"X" document of particular relevance; the considered movel or cannot be consider	e claimed invention cannot be red to involve an inventive step							
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other	document which may throw doubts on priority claim(s) or which is when the document is taken alone								
special reason (as specified)  "O" document referring to an oral disclosure, use, exhibition or other means	special reason (as specified)  "Y"  document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being								
"P" document published prior to the international filing date but later "A" document member of the same patent family than the priority date claimed									
Date of the actual completion of the international search	Date of mailing of the international se	earch report							
30 SEPTEMBER 2001									
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT	Authorized officer  JEHANNE SOUAYA  Telephone No. (708) 308-1835	1/Mili-							
Washington, D.C. 20231	Telephone No. /700\ 000 1005								

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
□ other:

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.